

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2002, 16:26:18 ; Search time 30 Seconds

(without alignments)
559.071 Million cell updates/sec

Title: US-09-868-953-1
Perfect score: 804

Sequence: 1 MFSTKALLGLISTALGP.....GIQSHVHIENCSSDPTAYL 151

Scoring table: BLOSUM62

Searched: 747574 seqs, 111073796 residues

total number of hits satisfying chosen parameters: 747574

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%

Database : A_Geneseq_032802:*

1:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1980.DAT *
2:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1981.DAT *
3:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1982.DAT *
4:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1983.DAT *
5:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1984.DAT *
6:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1985.DAT *
7:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1986.DAT *
8:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1987.DAT *
9:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1988.DAT *
10:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1989.DAT *
11:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1990.DAT *
12:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1991.DAT *
13:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1992.DAT *
14:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1993.DAT *
15:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1994.DAT *
16:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1995.DAT *
17:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1996.DAT *
18:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1997.DAT *
19:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1998.DAT *
20:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl1999.DAT *
21:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl2000.DAT *
22:	/SIDS1/gcgdata/genseq/genseqp-emb1/AAl2001.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	802	99.8	151	19	AAM52300	Human Mm-1 homoid
2	802	99.8	151	19	AAW53840	Human LECT2, Homo
3	802	99.8	151	21	AAB11932	Human leukocyte-de
4	792	98.5	151	17	AAR99271	Leukocyte cell-dere
5	734	91.3	681	22	ABG14781	Novel human diagno
6	698	86.8	151	19	AAW57630	Chondromodulin-ti
7	641	79.7	133	17	AAR97815	Human chondromodu
8	638	79.4	133	14	AAR32177	Bovine foetal cartil
9	294	36.6	54	22	ABB38145	Peptide #5651 enco
10	294	36.6	54	22	ABB38353	Protein #5352 enco
11	294	36.6	54	22	AAW58781	Human brain expres

[illegible]

ALIGNMENTS

RESULT	1
AAM52000	
ID	AAM52000 standard; Protein; 151 AA.
XX	
AC	AAM52000;
XX	
DT	15-MAR-2002 (first entry)
XX	
DE	Human Mim-1 homologous protein.
XX	
KW	Human; Mim-1 homologue.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Misc-difference 1 Location/Qualifiers
FT	/note= "encoded by ATO"
FT	Misc-difference 58
FT	/note= "encoded by ATO"
XX	
PN	KR97065715-A.
XX	
PD	13-OCT-1997.
XX	
PF	29-MAR-1996; 96KR-0009150.
XX	
PR	29-MAR-1996; 96KR-0009150.
XX	
PA	(ROAD) KOREA INST SCI & TECHNOLOGY.
PI	
PI	Choi IS, Chung TH, Kim JH, Lee YH, Lee IA, Choi YG, Song JC,
PI	Lhm JS;
XX	
XR	WPI; 1998-492656/42.

CC encompasses a method of screening candidate bone resorption inhibitors
 CC using LECT2, and the inhibitors thus identified. Human LECT2 was isolated
 CC from undifferentiated bone cells, and its ability to inhibit osteoclast-
 CC mediated bone resorption was examined via a pit assay. Human LECT2 caused
 CC 50% inhibition of bone resorption at 1 microgram/ml and 100% inhibition
 CC at 10 micrograms/ml. LECT2 or its derivatives may be used for the
 CC treatment of hypercalcaemia, osteoporosis, bone fractures, lumbago,
 CC (lumbodynia), hyperparathyroidism and Paget's disease.

XX Sequence 151 AA:

Query Match 99.8%; Score 802; DB 21; Length 151;
 Best Local Similarity 100.0%; Pred. No. 2,1e-87;
 Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MESTKALLAGLSTALAGPWANICAGKSSNEIRCDRHGCGQSAQRSPHGVXLC 60
 DB 1 MESTKALLAGLSTALAGPWANICAGKSSNEIRCDRHGCGQSAQRSPHGVXLC 60
 OY 61 SAGSTVYAPFTGMIVGEEKPYQNKNAINNGVRLSGRGCVMFYIKPIKKGPIKKEKL 120
 DB 61 SAGSTVYAPFTGMIVGEEKPYQNKNAINNGVRLSGRGCVMFYIKPIKKGPIKKEKL 120
 OY 121 GTLLPLOKVPYPGIOSHVHIENCSDSPPTAYL 151
 DB 121 GTLLPLOKVPYPGIOSHVHIENCSDSPPTAYL 151

RESULT 4

AAR99271
 ID AAR99271 standard; Protein: 151 AA.

XX AAR99271;

DT 13-DEC-1996 (first entry)

XX Leukocyte cell-derived chemotaxin LECT2b.

XX LECT2b; LECT2a; leukocyte cell-derived chemotaxin; cytokine;

KW neutrophil; chemotaxis; cancer; therapy; diagnosis.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 42 /note="residue 42 may also be Ile, owing to a
 FT polymorphism in the nucleotide sequence"

EP723016-A2.

24-JUL-1996.

27-NOV-1995; 95EP-0402670.

28-NOV-1994; 94JP-0293233.

(SUZU/) SUZUKI K.

Amiga S, Mizuno S, Suzuki K, Yamakawa Y, Yamagoe S;

WPI: 1996-335477/34.

N-PSDB; AAT35260.

XX New human leukocyte cell-derived chemotaxin cpds. - used for the
 PT detection and therapy of cancer and disorders of the cytokine
 PT network

PS Claim 2; Page 13; 17pp; English.

XX Novel human leukocyte cell-derived chemotaxin LECT2b (AAR99271) has
 CC neutrophil chemotactic activity. Its amino acid sequence was
 CC deduced from a cDNA clone (AAT35260) derived from
 CC phytohaemagglutinin-activated T-cell leukaemia cell line SKW-3.

CC Recombinant LECT2a can be produced, pref. as a fusion protein, in
 CC transformed host cells, esp. E. coli, yeast, insect cells, CHO,
 CC CV-1, 293, C127, 3T3, F-928, HeLa and SKW-3 cells. LECT2b and
 CC LECT2a (see also AAR9266-70) are useful for the diagnosis, therapy
 CC and prediction of cancer and disorders of the cytokine network.

XX Sequence 151 AA:

Query Match 98.5%; Score 792; DB 17; Length 151;
 Best Local Similarity 98.0%; Pred. No. 3.3e-86;
 Matches 148; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 MESTKALLAGLSTALAGPWANICAGKSSNEIRCDRHGCGQSAQRSPHGVXLC 60
 DB 1 MESTKALLAGLSTALAGPWANICAGKSSNEIRCDRHGCGQSAQRSPHGVXLC 60
 OY 61 SAGSTVYAPFTGMIVGEEKPYQNKNAINNGVRLSGRGCVMFYIKPIKKGPIKKEKL 120
 DB 61 SAGSTVYAPFTGMIVGEEKPYQNKNAINNGVRLSGRGCVMFYIKPIKKGPIKKEKL 120
 OY 121 GTLLPLOKVPYPGIOSHVHIENCSDSPPTAYL 151
 DB 121 GTLLPLOKVPYPGIOSHVHIENCSDSPPTAYL 151

RESULT 5

ABG14781
 ID ABG14781 standard; Protein: 681 AA.

XX ABG14781;

DT 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #14772.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

PN WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT;

WPI: 2001-639362/73.

N-PSDB; AAS78968.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

PS Claim 20; SEQ ID No 45140; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical

QY 19 GPMANICAGSSNEIRTCDRHGGCGQYSAQRSQRP HQGVDLCSAGSTVYAPFTGMIVGOE 78
 DB 1 GPMANICAGSSNEIRTCDRHGGCGQYTAQRNOKLHQGVDLCSAGSTVYAPFTGMIVGOE 60
 QY 79 KPYONKAINNGVIRISGGGFCIKMFYIKPKYKGSIKKGEKLTLLPLQKVPYPGIOSHHV 138
 DB 61 KPYONKAINNGVIRISGGGFCIKMFYIKPKYKGSIKKGEKLTLLPLQKVPYPGIOSHHV 120
 QY 139 IENCDSPTAYL 151
 DB 121 IENCDSLDPYVL 133
 RESULT 8
 AAR32177
 ID AAR32177 standard; protein; 133 AA.
 AC AAR32177;
 XX 04-JUN-1993 (first entry)
 DE Human foetal cartilage chondromodulin II.
 KW stimulating chondrocyte growth; fractures; treatment; cartilage;
 KM disorders.
 OS Homo sapiens.
 XX EP526883-A.
 PM 10-FEB-1993.
 PD 05-AUG-1992; 92EP-0113342.
 PE 05-AUG-1991; 91JP-0195495.
 PR 05-AUG-1991; 91JP-0195495.
 XX (MITU) MITSUBISHI KASEI CORP.
 PA Hiraki Y, Kamizono A, Kondo J, Suzuki F, Tanaka H;
 PI Teranishi Y;
 PI WPI: 1993-046832/06.
 DR New chondromodulin-II protein - stimulates chondrocyte growth,
 PT used for treating fractures and cartilage disorders
 PS Claim 1; Page 11; 18pp; English.
 CC This sequence represents human foetal cartilage chondromodulin II.
 CC It has a mol wt. of 16kD by SDS-PAGE, the ability to stimulate
 CC chondrocytes in the presence or absence of FGF, and the ability to
 CC promote differential potency of chondrocytes. The protein was
 CC fractionated and purified on the basis of the uptake of thymidine
 CC into cartilage cells. The sequence of the full protein was obtained
 CC by aligning N-terminal sequences and internal sequences. Tests
 CC revealed that the protein possessed a potent stimulating effect on
 CC chondrocyte growth ie. uptake of radioactive thymidine in presence
 CC of 1200ng/ml chondromodulin was increased 3.5 fold of that observed
 CC in the absence of FGF and 1.4 fold of that observed in the presence
 CC of 0.4ng FGF.
 CC
 SQ Sequence 133 AA;
 Query Match 79.4%; Score 638; DB 14; Length 133;
 Best Local Similarity 87.2%; Pred. No. 6.3e-68;
 Matches 116; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
 QY 19 GPMANICAGSSNEIRTCDRHGGCGQYSAQRSQRP HQGVDLCSAGSTVYAPFTGMIVGOE 78
 DB 1 GPMANICAGSSNEIRTCDRHGGCGQYTAQRNOKLHQGVDLCSAGSTVYAPFTGMIVGOE 60
 QY 79 KPYONKAINNGVIRISGGGFCIKMFYIKPKYKGSIKKGEKLTLLPLQKVPYPGIOSHHV 138

DB 61 KPYONKAINNGVIRISGGGFCIKMFYIKPKYKGSIKKGEKLTLLPLQKVPYPGIOSHHV 120
 QY 139 IENCDSPTAYL 151
 DB 121 IENCDSLDPYVL 133
 RESULT 9
 ABB38145
 ID ABB38145 standard; Peptide; 54 AA.
 AC ABB38145;
 XX 04-FEB-2002 (first entry)
 DE Peptide #5651 encoded by human foetal liver single exon probe.
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 OS Homo sapiens.
 XX WO200157277-A2.
 PN 09-AUG-2001.
 PD 30-JAN-2001; 2001MO-US00669.
 PF 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI: 2001-483447/52.
 DR Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 PS Claim 27; SEQ ID NO 30780; 639pp + sequence listing; English.
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences.
 CC
 SQ Sequence 54 AA;
 Query Match 36.6%; Score 294; DB 22; Length 54;
 Best Local Similarity 100.0%; Pred. No. 1.9e-27;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 98 FCVKKMFYIKPKYKPKYKGEKLTLLPLQKVPYPGIOSHHV IENCDSPTAYL 151
 DB 1 FCVKKMFYIKPKYKPKYKGEKLTLLPLQKVPYPGIOSHHV IENCDSPTAYL 54
 RESULT 10
 ABB23353
 ID ABB23353 standard; Protein; 54 AA.
 AC ABB23353;

```

XX 23-JAN-2002 (first entry)
XX Protein #5352 encoded by probe for measuring heart cell gene expression.
DE Human: gene expression: heart; microarray; vascular system;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease.
XX Homo sapiens.
XX WO200157274-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00666.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX Claim 15; SEQ ID NO 25123; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX ABA21535-ABA41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 54 AA:
XX
XX Query Match 36.6%; Score 294; DB 22; Length 54;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-27;
XX Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 98 FCVAMFYIKPIKYGPIKKGKGLTLLPLQKYVPGIQSHVHTENCSSDPTAYL 151
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 1 FCVAMFYIKPIKYGPIKKGKGLTLLPLQKYVPGIQSHVHTENCSSDPTAYL 54
XX
XX RESULT 11
XX ID AAM58781 standard; Protein: 54 AA.
XX AC AAM58781.
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30886.
XX KW Human: brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer.

```

```

XX Homo sapiens.
XX WO200157275-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00667.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX Example 4; SEQ ID NO: 30886; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.
XX
XX Sequence 54 AA:
XX
XX Query Match 36.6%; Score 294; DB 22; Length 54;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-27;
XX Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 98 FCVAMFYIKPIKYGPIKKGKGLTLLPLQKYVPGIQSHVHTENCSSDPTAYL 151
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 1 FCVAMFYIKPIKYGPIKKGKGLTLLPLQKYVPGIQSHVHTENCSSDPTAYL 54
XX
XX RESULT 12
XX ID AAM71294 standard; Protein: 54 AA.
XX AC AAM71294.
XX DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31600.
XX KW Human: bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukemia; lymphoma; myeloma.
XX Homo sapiens.
XX WO200157276-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00668.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.

```

PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 31600; 658bp + Sequence listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is a
XX protein encoded by one of the probes of the invention.
XX
XX Sequence 54 AA;
XX

Query Match 36.6%; Score 294; DB 22; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.9e-27;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 FCVKMFYIKPIKYGPIKKGKGLTLPLQKYVPGIQSHVHINCSDPTAYL 151
DB 1 FCVKMFYIKPIKYGPIKKGKGLTLPLQKYVPGIQSHVHINCSDPTAYL 54

RESULT 13

AAM18972
ID AAM18972 standard; Protein: 54 AA.

XX AAM18972;

XX 12-OCT-2001 (first entry)

XX Peptide #5406 encoded by probe for measuring cervical gene expression.

XX Probe; human; microarray; gene expression; cervical epithelial cell;

XX cervical cancer.

XX Homo sapiens.

XX WO200157278-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US006670.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 27; SEQ ID No 23798; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes

CC (SENP: see AAI10068-AA128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX

XX Sequence 54 AA;

Query Match 36.6%; Score 294; DB 22; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.9e-27;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 FCVKMFYIKPIKYGPIKKGKGLTLPLQKYVPGIQSHVHINCSDPTAYL 151
DB 1 FCVKMFYIKPIKYGPIKKGKGLTLPLQKYVPGIQSHVHINCSDPTAYL 54

RESULT 14

AAM31573
ID AAM31573 standard; Protein: 54 AA.

XX AAM31573;

XX 17-OCT-2001 (first entry)

XX Peptide #5610 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;

XX genetic disorder.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-48897/53.

XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
XX Claim 27; SEQ ID No 31842; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
XX see AAI13135-AA157546). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders.
XX

XX Sequence 54 AA;

Query Match 36.6%; Score 294; DB 22; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.9e-27;

Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 FCVKMFYIPKPKKKEKGLTLLPQKVPKIOSHVHIENCSSDPTAYL 151
 Db 1 FCVKMFYIPKPKKKEKGLTLLPQKVPKIOSHVHIENCSSDPTAYL 54

RESULT 15

AA99266
 ID AAR99266 standard; Peptide; 54 AA.

AA99266;

13-DEC-1996 (first entry)

Leukocyte cell-derived chemotaxin LECT2a peptide 1.

LECT2a; LECT2b; leukocyte cell-derived chemotaxin; cytokine; neutrophil; chemotaxis; cancer; therapy; diagnosis.

Homo sapiens.

Key Location/Qualifiers

Misc-difference 53 /label= unidentified amino acid

EP723016-A2.

24-JUL-1996.

27-NOV-1995; 95EP-0402670.

28-NOV-1994; 94JP-0293233.

(SUZUKI) SUZUKI K.

Amiga S, Mizuno S, Suzuki K, Yamakawa Y, Yamagoe S;

WPI; 1996-335477/34.

New human leukocyte cell-derived chemotaxin cpds. - used for the detection and therapy of cancer and disorders of the cytokine network

Claim 1; Page 11; 17pp; English.

LECT2a peptides 1-5 (AAR99266-70) comprise tryptic peptides of novel human leukocyte cell-derived chemotaxin LECT2a, in order from the N-terminal sequence. They were obtd. by tryptic digestion of 15 kDa LECT2a isolated from the phytohemagglutinin-activated T-cell leukemia cell line SKW-3. LECT2a and LECT2b (AAR99271) have neutrophil chemotactic activity. They are useful for the diagnosis, therapy and prediction of cancer and disorders of the cytokine network.

Sequence 54 AA;

Query Match 31.3%; Score 252; DB 17; Length 54;

Best Local Similarity 83.3%; Pred. NO. 1.9e-22; Mismatches 45; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 19 GPMANICAGKSSNEIRTCDRHCGQYSAORSQRP HQGVDXLCSAGSTVAPFTG 72
 Db 1 GPMANICAGKSSNEIRTCDRHCGQYSAORSQRP HQGVDXLCSAGSTVAPFTG 54

Search completed: October 21, 2002, 16:33:01
 Job time : 32 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2002, 16:32:28 ; Search time 13 Seconds

(without alignments).
283.713 Million cell updates/sec

Title: US-09-868-953-1

Perfect score: 804

Sequence: 1 MESTKALLAGLSTALGP.....GIQSHVHENCSDPPAYL 151

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Print. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	802	99.8	151	2	US-08-563-148E-6
2	802	99.8	151	4	US-09-194-139-1
3	802	99.8	151	4	US-09-194-139-1
4	417	79.7	133	1	US-07-924-753-15
5	252	31.3	54	2	US-09-194-139-9
6	192	23.9	41	2	US-08-563-148E-1
7	132	16.4	27	1	US-07-924-753-1
8	95	11.8	20	2	US-08-563-148E-5
9	94	11.7	18	1	US-07-924-753-4
10	89	11.1	19	1	US-07-924-753-12
11	86	10.7	18	1	US-07-924-753-11
12	86	10.7	19	1	US-07-924-753-3
13	74	9.2	702	6	5240838-15
14	73	9.1	14	1	US-07-924-753-10
15	69	8.6	14	1	US-07-924-753-9
16	67.5	8.4	625	3	US-08-996-139-15
17	67.5	8.4	625	4	US-08-995-659-15
18	67.5	8.4	625	4	US-08-995-659-15
19	66.5	8.3	2285	4	US-09-215-649A-15
20	65.5	8.1	1241	4	US-09-040-774-2
21	64	8.0	428	3	US-09-118-319-5
22	63.5	7.9	2595	4	US-09-036-987A-2
23	63.5	7.9	2595	4	US-09-370-700-2
24	63	7.8	14	4	US-09-194-139-10
25	63	7.8	2004	1	US-08-375-709-15
26	63	7.8	2004	1	US-08-752-929-15
27	63	7.8	2004	4	US-09-090-793-9

28	62.5	7.8	406	1	US-08-293-778-24
29	62.5	7.8	406	1	US-08-295-411-5
30	62.5	7.8	406	2	US-08-955-471-5
31	62.5	7.8	406	5	PCP-US92-10242-5
32	62.5	7.8	433	3	US-08-926-842B-61
33	62.5	7.8	444	2	US-08-475-845-2
34	62.5	7.8	444	2	US-08-327-690-2
35	62.5	7.8	444	2	US-08-660-289-2
36	62.5	7.8	444	2	US-08-537-807-2
37	62.5	7.8	444	2	US-08-871-003-2
38	62.5	7.8	444	3	US-08-464-233-2
39	62.5	7.8	444	4	US-09-189-607-2
40	62.5	7.8	444	4	US-09-378-907-2
41	62.5	7.8	444	5	PCP-US94-05779-2
42	62.5	7.8	466	1	US-07-882-202A-4
43	62.5	7.8	466	1	US-08-021-615A-4
44	62.5	7.8	466	1	US-08-321-777-4
45	62.5	7.8	466	4	US-09-009-217-14

ALIGNMENTS

RESULT 1
US-08-563-148E-6
; Sequence 6, Application US/08563148E
; Patent No. 5929224
; GENERAL INFORMATION:
; APPLICANT: Kazuo SUZUKI et al.
; TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESS: 663-2, Shligi
; STREET: Misaki-machi
; CITY: Isumi-gun
; STATE: Chiba-ken
; COUNTRY: JAPAN
; ZIP: 299-45
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; OPERATING SYSTEM: IBM compatible
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/563, 148E
; FILING DATE: No. 5929224ember 27, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 293233/1994
; FILING DATE: 28-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5220
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single strand
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: intermediate fragment
; ORIGINAL SOURCE:
; ORGANISM: human
; TISSUE TYPE: T-cell derived leukemia cells
; PUBLICATION INFORMATION:
; AUTHORS: Kazuo SUZUKI et al.
; TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 151
US-08-563-148E-6

Query Match 99.8%; Score 802; DB 2; Length 151;
Best Local Similarity 99.3%; Pred. No. 8.9e-90;
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MESTKALLAGLSTALAGPMANICAGKSSNEIRTCDRHGGCGYSAORSQRPBGVDXLC 60
DB 1 MESTKALLAGLSTALAGPMANICAGKSSNEIRTCDRHGGCGYSAORSQRPBGVDXLC 60

OY 61 SAGSTVYAPFTGMIVGGEKPYQNNAINNGVIRISGRGFCVMMFYIKPIKYGPIKKEKL 120
DB 61 SAGSTVYAPFTGMIVGGEKPYQNNAINNGVIRISGRGFCVMMFYIKPIKYGPIKKEKL 120

OY 121 GTLLPLOKVPYGIOSHVIHENCSSDPTAYL 151
DB 121 GTLLPLOKVPYGIOSHVIHENCSSDPTAYL 151

RESULT 2

US-09-194-139-1
Sequence 1, Application US/09194139A

Patent No. 6306608

GENERAL INFORMATION:

APPLICANT: Arai, Takao

TITLE OF INVENTION: ANTI-HUMAN LECT2 ANTIBODY, CELLS PRODUCING THE SAME,

TITLE OF INVENTION: AND METHOD AND KIT FOR ASSAYING THE SAME

FILE REFERENCE: US Application 09/194,139

CURRENT APPLICATION NUMBER: US/09/194,139A

CURRENT FILING DATE: 1998-11-25

EARLIER APPLICATION NUMBER: JP 8-132160

EARLIER FILING DATE: 1996-05-27

NUMBER OF SEQ ID NOS: 10

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO. 1

LENGTH: 151

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: residue 58 is either Valine or Isoleucine

US-09-194-139-1

Query Match 99.8%; Score 802; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 8.9e-90;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MESTKALLAGLSTALAGPMANICAGKSSNEIRTCDRHGGCGYSAORSQRPBGVDXLC 60
DB 1 MESTKALLAGLSTALAGPMANICAGKSSNEIRTCDRHGGCGYSAORSQRPBGVDXLC 60

OY 61 SAGSTVYAPFTGMIVGGEKPYQNNAINNGVIRISGRGFCVMMFYIKPIKYGPIKKEKL 120
DB 61 SAGSTVYAPFTGMIVGGEKPYQNNAINNGVIRISGRGFCVMMFYIKPIKYGPIKKEKL 120

OY 121 GTLLPLOKVPYGIOSHVIHENCSSDPTAYL 151
DB 121 GTLLPLOKVPYGIOSHVIHENCSSDPTAYL 151

RESULT 3

US-07-924-753-15
Sequence 15, Application US/07924753

Patent No. 5270303

GENERAL INFORMATION:

APPLICANT: Fujio SUZUKI et al.

TITLE OF INVENTION: CHONROMODULIN-II PROTEIN

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Wenderoth, Lind & Ponack

STREET: 805 Fifteenth Street, N.W., #700

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/924,753

FILING DATE: 19920804

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

AUTHOR/AGENT INFORMATION:

NAME: Warren M. Cheek, Jr.

REGISTRATION NUMBER: 33,367

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-8850

TELEFAX:

TELEX:

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 133 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM: bovine

STRAIN:

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE: fetal cartilage

CELL TYPE:

ORGANELLE:

IMMEDIATE SOURCE:

LIBRARY:

CLONE:

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATE:

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO:

US-07-924-753-15

Query Match 79.7%; Score 641; DB 1; Length 133;
Best Local Similarity 87.2%; Pred. No. 2.7e-70;
Matches 116; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

OY 19 GPWANICAGKSSNEIRTCDRHGGCGYSAORSQRPBGVDXLCAGSTVYAPFTGMIVGGE 78
DB 1 GPWANICAGKSSNEIRTCDRHGGCGYTAORXNOKLHOGVDXCDGSGTVYAPFTGKINGOE 60

QY 79 KPYONKNAINGVRISGSGFCYKMFYIKPIKKGKLTLLPLQKVPYPGIOSHVH 138
DB 61 KPYONKNAINGVRISGSGFCYKMFYIKPIKKGKLTLLPLQKVPYPGIOSHVH 120
QY 139 IENCSSDPTAYL 151
DB 121 IENCDSLDPYVL 133

RESULT 4

US-09-194-139-9
Sequence 9, Application US/09194139A
Patent No. 6306608
GENERAL INFORMATION:
APPLICANT: Arai, Takao
TITLE OF INVENTION: ANTI-HUMAN LECT2 ANTIBODY, CELLS PRODUCING THE SAME,
TITLE OF INVENTION: AND METHOD AND KIT FOR ASSAYING THE SAME
FILE REFERENCE: US application 09/194,139
CURRENT FILING DATE: 1998-11-25
EARLIER APPLICATION NUMBER: JP 8-132160
EARLIER FILING DATE: 1996-05-27
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 9
LENGTH: 98
TYPE: PRT
ORGANISM: Bovine
US-09-194-139-9

Query Match
Best Local Similarity 51.9%; Score 417; DB 4; Length 98;
Matches 84; Conservative 6; Mismatches 8; Indels 32; Gaps 4;

QY 19 GPMANICAGKSSNEIRTCDRHCGGQYSAQRSOPRHOGVDXLCGASGTVYAPFTGMIYGOE 78
DB 1 GPMANICAGKSSNEIRTCDRHCGGQYSAQRSOPRHOGVDXLCGASGTVYAPFTGMIYGOE 59
QY 79 KPYONKNAINGVRISGSGFCYKMFYIKPIKKGKLTLLPLQKVPYPGIOSHVH 138
DB 60 KPYKN-----ISGGGFC-----IKYKGS-----VYRGIQSHIH 88
QY 139 IENCSSDPT 148
DB 89 IENCDSLDP 98

RESULT 5

US-08-563-148E-1
Sequence 1, Application US/08563148E
Patent No. 5929224

GENERAL INFORMATION:
APPLICANT: Kazuo SUZUKI et al.
TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: 663-2, Shiigi
STREET: Misaki-machi
CITY: Tsumi-gun
STATE: Chiba-ken
COUNTRY: JAPAN
ZIP: 299-45
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS V.5
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/563,148E
FILING DATE: NO. 5929224ember 27, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 293233/1994
FILING DATE: 28-NOV-1994

ATTORNEY/AGENT INFORMATION:
NAME: C. Bruce Hamburg
REGISTRATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-5230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)986-2340
TELEFAX: (212)953-7733
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
STRANDEDNESS: single strand
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal fragment
ORIGINAL SOURCE:
ORGANISM: human
TISSUE TYPE: T-cell derived leukemia cells
PUBLICATION INFORMATION:
AUTHORS: Kazuo SUZUKI et al.
TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 to 54
US-08-563-148E-1

Query Match
Best Local Similarity 31.3%; Score 252; DB 2; Length 54;
Matches 45; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 19 GPMANICAGKSSNEIRTCDRHCGGQYSAQRSOPRHOGVDXLCGASGTVYAPFTG 72
DB 1 GPMANICAGKSSNEIRTCDRHCGGQYSAQRSOPRHOGVDXLCGASGTVYAPFTG 54

RESULT 6

US-07-924-753-1
Sequence 1, Application US/07924753
Patent No. 5270303

GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Displaywrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: N-terminal fragment
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-1

Query Match 23.9%; Score 192; DB 1; Length 41;
Best Local Similarity 82.9%; Pred. No. 2e-16;
Matches 34; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 19 GPMATCAGKSSNEIRTCDRHGGCCGYSGQRSPHQVDXL 59
DB 1 GPMATCAGKSSNEIRTCDRHGGCCGYSGQRSPHQVDXL 41

RESULT 7
US-07-924-753-14
Sequence 14, Application US/07924753
Patent No. 5270303
GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE//DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: internal fragment
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-14

Query Match 16.4%; Score 132; DB 1; Length 27;
Best Local Similarity 92.6%; Pred. No. 2.2e-09;
Matches 25; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 111 KSPIKGEKGLTLLPLQKVPPIQSHV 137
DB 1 KSPIKGEKGLTLLPLQKVPPIQSHI 27

RESULT 8
US-08-563-148E-5
Sequence 5, Application US/08563148E
Patent No. 5929224
GENERAL INFORMATION:
APPLICANT: Kazuo SUZUKI et al.

TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: 663-2, Shligi
STREET: Misaki-machi
CITY: Isumi-gun
STATE: Chiba-ken
COUNTRY: JAPAN
ZIP: 299-45
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS V.5
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/563,148E
FILING DATE: NO. 5929224ember 27, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 293233/1994
FILING DATE: 28-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: C. Bruce Hamburg
REGISTRATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-5230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)986-2340
TELEFAX: (212)953-7733
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single strand
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: intermediate fragment
ORIGINAL SOURCE:
ORGANISM: human
TISSUE TYPE: T-cell derived leukemia cells
PUBLICATION INFORMATION:
AUTHORS: Kazuo SUZUKI et al.
TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 to 20
US-08-563-148E-5

Query Match 11.8%; Score 95; DB 2; Length 20;
Best Local Similarity 85.0%; Pred. No. 4.7e-05;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

129 VYPGISHVHIENCSSDPT 148
|||||:|||||
Db 1 VYPGISHIHIENCLSDPT 20

RESULT 9
US-07-924-753-4
; Sequence 4, Application US/07924753
; Patent No. 5270303
; GENERAL INFORMATION:
; APPLICANT: Fujio SUZUKI et al.
; TITLE OF INVENTION: CHONDRODULIN-II PROTEIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: internal fragment
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-4

Query Match 11.7%; Score 94; DB 1; Length 18;
Best Local Similarity 88.9%; Pred. No. 5.4e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

0Y 129 VYPGISHVHIENCSSD 146
|||||:|||||
Db 1 VYPGISHIHIENCLSD 18

RESULT 10
US-07-924-753-12
; Sequence 12, Application US/07924753

Patent No. 5270303
GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: Internal fragment
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:

FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-12
Query Match 11.1%; Score 89; DB 1; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00023;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 68 APTGMIYQEKRYONKNA 86
|||||:|||||:|||||
Db 1 APTGKIMGQEKRYKNKNA 19
RESULT 11
US-07-924-753-11
Sequence 11, Application US/07924753
Patent No. 5270303
GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: Internal fragment
ORIGINAL SOURCE:
ORGANISM: bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:

Db 525 GEFQKAVELPFSSILS-LSNNEVL-----QYLASRAQRNRNAGYLLLEDAENAEVQI 575
QY 71 TGMIVGQEKPYQKNAINNVRISRGF-----CYKMEYIKIKKGP1--KKEGKL 120
Db 576 IG-VGAEMEFADKAA-----KILGRKFRFRVLSPCTRLEPDSIGYFRSRLRKDGROY 628
QY 121 GLLPLQKVPYPGIQSHV 137
Db 629 PTVV-----VDAAHV 637

RESULT 14
US-07-924-753-10
Sequence 10, Application US/07924753
Patent No. 5270303
GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: internal fragment
ORIGINAL SOURCE:
ORGANISM: Bovine
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: fetal cartilage
CELL TYPE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:

UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-10

Query Match 9.1%; Score 73; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 KGLTLPLQKVPYPG 132
Db 1 KGLTLPLQKVPYPG 14

RESULT 15
US-07-924-753-9
Sequence 9, Application US/07924753
Patent No. 5270303
GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,753
FILING DATE: 19920804
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:


```

; FRAGMENT TYPE: internal fragment
; ORIGINAL SOURCE:
; ORGANISM: bovine
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE: fetal cartilage
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-924-753-9

Query Match      8.6%; Score 69; DB 1; Length 14;
Best Local Similarity 92.9%; Pred. No. 0.042;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 103 FYIKPIKYKGPICK 116
   |||||
Db 1 FYIKPIKYKGSIKK 14

```

Search completed: October 21, 2002, 16:34:37
 Job time : 14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2002, 16:31:03 : Search time 17 seconds

(without alignments)
853.499 Million cell updates/sec

Title: US-09-868-953-1

Perfect score: 804

Sequence: 1 MFSTKALLAGLISTALGP.....GIQSHVHTENCSSDPTAYL 151

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	641	79.7	133	2 JH0270	chondromodulin II
2	429	53.4	326	2 A33755	myb-induced myeloid
3	99.5	12.4	305	2 T16567	hypothetical prote
4	80.5	10.0	417	2 C62840	conserved hypothet
5	76.5	9.5	475	2 F64151	hypothetical prote
6	76	9.5	240	2 H84197	hypothetical prote
7	76	9.5	651	2 T42644	hypothetical prote
8	74	9.2	750	2 F90571	lipoprotein [impor
9	73	9.1	350	2 AD0466	hypothetical prote
10	73	9.1	710	1 XJH0KX	formaldehyde trans
11	72.5	9.0	428	2 T36365	probable penicilli
12	71.5	8.9	295	2 AC2451	hypothetical prote
13	71.5	8.9	706	2 T01351	subtilisin-like pr
14	71	8.8	284	2 E97836	probable periplasm
15	70	8.7	444	2 T05614	hypothetical prote
16	69.5	8.6	268	2 D97054	hypothetical prote
17	69	8.6	1331	1 XORTDH	probable membrane
18	69	8.6	1434	2 T32647	xanthine dehydroge
19	68	8.5	184	2 T15360	hypothetical prote
20	68	8.5	310	2 A60967	hypothetical prote
21	68	8.5	762	2 D70407	insulin-like growt
22	68	8.5	3010	1 S18030	DNA mismatch repla
23	67.5	8.4	273	2 E81330	genome polyprotein
24	67.5	8.4	312	2 AE2215	probable periplasm
25	67.5	8.4	355	2 T51414	hypothetical prote
26	67.5	8.4	535	2 C36811	CONSTANS-like 1 -
27	67.5	8.4	535	2 A37994	hypothetical prote
28	67	8.3	398	2 A35281	Rfl protein - salm
29	67	8.3	554	2 C98133	integraly muc

30	67	8.3	554	2 AE3154	agrobacterium vlrn
31	67	8.3	689	2 G84447	hypothetical prote
32	67	8.3	1052	2 T14343	zinc finger RNA bi
33	67	8.3	1347	2 T45632	helicase homolog F
34	66.5	8.3	2285	2 T12796	probable transglyc
35	66	8.2	320	2 C85440	myb-related protei
36	66	8.2	447	2 B83563	conserved hypothet
37	66	8.2	782	2 S18032	genome polyprotein
38	65.5	8.1	233	2 AD3344	hypothetical prote
39	65.5	8.1	364	2 T12039	cysteine proteinas
40	65.5	8.1	364	2 T46630	cysteine proteinas
41	65.5	8.1	454	2 AE3537	formyl-coenzyme a
42	65.5	8.1	606	2 H97678	hypothetical prote
43	65.5	8.1	606	2 AP2903	GTP-binding tyrosi
44	65.5	8.1	608	2 AB3562	GTP-binding protei
45	65.5	8.1	621	2 D96554	hypothetical prote

ALIGNMENTS

RESULT 1

JH0270

chondromodulin II - bovine

N:Alternate names: cartilage-derived factor

C:Species: Bos primigenius taurus (cattle)

C:Date: 07-Jul-1997 #sequence revision 18-Jul-1997 #text change 17-Mar-1999

C:Accession: JH0270

R:Hitaki, Y.; Inoue, H.; Kondo, J.; Kamizono, A.; Yoshitake, Y.; Shukunam, C.;

J. Biol. Chem. 271, 22657-22662, 1996

A:Title: A novel growth-promoting factor derived from fetal bovine cartilage, ch

A:Reference number: JH0270; MUID:96394331

A:Accession: JH0270

A:Molecule type: protein

A:Residues: 1-133 <H1R>

A:Experimental source: cartilage cell

C:Comment: This protein stimulates proteoglycan synthesis in the growth plate ch

Query Match 79.7% Score 641; DB 2; Length 133;
Best Local Similarity 87.2% Pred. No. 6.5e-59;
Matches 116; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY	19	GPVANIACGSSNEIRCDHGGCGGYSAORSRPHOGVDPXICGASGVYVAFPTGMVIOE 78
DB	1	GPVANIACGSSNEIRCDHGGCGGYTQORNKLRHGVVLCDSGVVAFPTGKIMGQE 60
QY	79	KPYONKNAINNGVRIISGRCGVKMFYIKPIKYGKPIKGEKLTLLPLQKVYVGIOQSHV 138
DB	61	KPYKKNKAINNGVRIISGRCGVKMFYIKPIKYGKSIKGEKLTLLPLQKVYVGIOQSHI 120
QY	139	IENCDSDDPTAYL 151
DB	121	IENCDSDDPTAYL 133

RESULT 2

A33755

myb-induced myeloid protein 1 (min-1) precursor - chicken

N:Alternate names: arginine-specific ADP-ribosyltransferase endogenous substrate

C:Species: Gallus gallus (chicken)

C:Date: 30-Mar-1990 #sequence revision 30-Mar-1990 #text change 04-Sep-1998

C:Accession: A33755; S29229

R:Ness, S.A.; Marknell, A.; Graf, T.

A:Title: The v-myb oncogene product binds to and activates the promyelocyte-speci

A:Reference number: A33755; MUID:90090611

A:Accession: A33755

A:Molecule type: mRNA

A:Residues: 1-326 <MES>

A:Cross-references: GB:M29449; NID:g212341; PID:g212342

R:Yamada, K.; Tsuchiya, M.; Mishima, K.; Shimoyama, M.

FEBS Lett. 311, 203-205, 1992

A:Title: p33, an endogenous target protein for arginine-specific ADP-ribosyltrans

A:Reference number: S29229; MUID:93011981
 A:Accession: S29229
 A:Molecule type: protein
 A:Residues: 24-46;69-100;116-130;143-166;200-213;219-319 <YAM>
 C:Comment: This protein is a secretable component of promyelocyte granules. It is also a
 C:Genetics:
 A:Gene: m1m-1
 A:Note: this gene is activated by the product of the v-myb gene
 C:Keywords: duplication
 F:24-326/Product: myb-induced myeloid protein 1 #status experimental <MAT>
 Query Match 53.4%; Score 429; DB 2; Length 326;
 Best Local Similarity 50.3%; Pred. No. 1.3e-36;
 Matches 80; Conservative 26; Mismatches 39; Indels 14; Gaps 2;
 7 LLLAGTALAGP-----WANTCAGKSNIEPTCDRHGCCGYASQRSQR--P 52
 7 IALLSLVSTAFARQWEVHPPOQGRHMAQICSGNPFRRRCDRYCGNGYASRGKGEK 66
 53 HGVGDXLCSAGSTVYAPFTGMIVGQEKPYONKNAINNGVRIISRGFCVKMFYIKIKYK 112
 67 HKGVDTICTDGSIVYAPFSQSLSPRIFFHNGNAIDGVOIGSGCYKVLICHPRIKYHG 126
 OY 113 PIKKGKLTLLPLOKVPBGIOSVHENDSDSPFAYL 151
 127 QIQKGQQLRMPLPMQKVPFGIVSHIHENCQSDPTHL 165
 Db
 RESULT 3
 T16567
 hypothetical protein K05F1.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16567
 R:Wohlmann, P.
 submitted to the EMBL Data Library, June 1995
 A:Description: The sequence of C. elegans cosmid K05F1.
 A:Reference number: Z18537
 A:Accession: T16567
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-305 <WOH>
 A:Cross-references: EMBL:U9377; NID:9868173; PID:9868182; PIDN:AAA68719.1; CESP:K05F1.4
 A:Experimental source: strain Bristol N2
 C:Genetics:
 A:Gene: CESP:K05F1.4
 A:introns: 44/2; 100/1; 164/1; 188/2; 295/1
 Query Match 12.4%; Score 99.5; DB 2; Length 305;
 Best Local Similarity 25.4%; Pred. No. 0.014;
 Matches 29; Conservative 21; Mismatches 37; Indels 27; Gaps 5;
 OY 55 GVDXLCASGTYVAPFTGMIVGQEKPY--ONKNAINNGVRIISRGFCVKMFYIKP 107
 52 GVDVRCHELGPPIYAPIGEMVFW--RPYGGKREKSCADGVRTGVMGQYAVHISVKL 110
 Db
 OY 108 IRKKGPIKKGKICTLPL-----QKVPYGIOASHV-----HIENC 142
 111 SFGGHEADDEIGELNRYCFNDRGQNDVPEHVEIRLYKESRLIDPTHTLQNC 164
 Db
 RESULT 4
 C82840
 conserved hypothetical protein Xf0167 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C:Accession: C82840
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717
 A:Note: for a complete list of authors see reference number A59328 below

A:Accession: C82840
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-417 <SIM>
 A:Cross-references: GB:AE003870; GB:AE003849; NID:99104955; PIDN:AAE82980.1; GSPDB
 A:Experimental source: strain 9a5c
 R:Simson, A.J.G.; Reineck, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga,
 Britones, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carriaro, D.M.; Carri
 as-Nevo, E.; Docena, C.; El-Dorty, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to Genbank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferrio, J.A.; Fraga, J.S.; Franco, M.C.;
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.;
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Mart
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki,
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sa
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da S
 M.; Teshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0167
 Query Match 10.0%; Score 80.5; DB 2; Length 417;
 Best Local Similarity 25.0%; Pred. No. 1.8;
 Matches 29; Conservative 24; Mismatches 42; Indels 21; Gaps 7;
 OY 34 RTCDRHCCGCGYASRSQRPHGVGDXLCSAGSTVYVAPFTGMIVGQEKPYONKNAINNGV 91
 Db 249 RLSTFGLRRHPILGTMMHKGVDAAATGPIFMAAGARVFGIQGRKGVAVILNH-- 306
 OY 92 RISGRFCVVMFYIKPIKPIKKGKEL--GTLPLQKVPBGIOS----HVHIE 140
 Db 307 ---GKGV--RLTYGHMSRF--GKIKAGQKINGTIVG---YVGMGLATGPHLHYE 352
 RESULT 5
 F64151
 hypothetical protein H10409 - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
 C:Accession: F64151
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlav
 R.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weid
 D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Ve
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630
 A:Accession: F64151
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-475 <TRIG>
 A:Cross-references: GB:U32724; GB:I42023; NID:91573378; PIDN:AAC22068.1; PID:91573
 A:Note: best homolog was a hypothetical protein from Escherichia coli
 C:Superfamily: hypothetical protein H10409
 Query Match 9.5%; Score 76.5; DB 2; Length 475;
 Best Local Similarity 28.7%; Pred. No. 5.3;
 Matches 27; Conservative 15; Mismatches 41; Indels 11; Gaps 4;
 OY 51 RPHGVGDXLCSAGSTVYVAPFTGMIVGQEKPYONKNAINNGVRIISRGFCVKMFYIKP--I 108
 Db 346 RPHGVGDFSVSQGPVIAPADGTV--EKVAYAGAGGRYVLMRHGRBYOTVYMLSKSLV 403
 OY 109 KYKGPICKKGIKTLPLQKVPBGIOS--HVHIE 140
 Db 404 KAGQTVKKGKRIAL-----SGNTGISTGPHLHYE 432
 RESULT 6
 H84197
 hypothetical protein Vng0391c [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence,revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: H84197
 R:Ng, W.V.; Kennedy, S.P.; Mahajiras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leihauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo Juny, K.H.; Alam, M.; Freilias, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
 A:Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: A84160; MUID:20504483
 A:Accession: H84197
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-240 <STO>
 A:Cross-references: GB:AE004437; NID:g10580006; PIDN:AGI8948.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: YNG0391C
 C:Superfamily: Hypothetical protein M0570

Query Match 9.5%; Score 76; DB 2; Length 240;
 Best Local Similarity 23.4%; Pred. No. 3;
 Matches 37; Conservative 23; Mismatches 56; Indels 42; Gaps 8;
 QY 9 LALISTALGPMANICAGKSSNEIRT-----CDRHGGGYSASORSQPHOGVDXCSA 62
 DB 98 VADTLDGLAG---VIAGAVESDFOTSMEGMDRLGCELYAPLWQPPRELADAMLTA 153
 QY 63 GSVVYAPFTGMIV-----GQKRPYONKN-----AINN--GYRISRGCVKMFYI 105
 DB 154 G-----FETRIIOVAAYVGLDKSWLGRRLDRDALAALALNDYVHMGEGGEFTLV 207
 QY 106 KPIKYGPIKKEKGLTLPLQKVPYGIOSVHI-ENC 142
 DB 208 DAPHMDRPIR-----LDYETWAGDRGHIEITEAC 237

RESULT 7
 T42644
 Hypothetical protein DKFZp566N1047.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 11-Jan-2000 #sequence,revision 11-Jan-2000 #text_change 11-Jan-2000
 C:Accession: T42644
 R:Ottewill, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, November 1999
 A:Reference number: Z22331
 A:Accession: T42644
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-651 <AAA>
 A:Cross-references: EMBL:AL133109
 A:Experimental source: fetal kidney; clone DKFZp566N1047
 C:Genetics:
 A:Note: DKFZp566N1047.1

Query Match 9.5%; Score 76; DB 2; Length 651;
 Best Local Similarity 31.4%; Pred. No. 8.2;
 Matches 27; Conservative 14; Mismatches 25; Indels 20; Gaps 5;
 QY 43 OXSORSORHGVDCXCSAGSTYAP-----FTGMTVGEKPYONKNAINN--GVRI 93
 DB 500 QINAQSSQMPQ-----AAQQAQIOPVLSGQOGFGLI-GVQOPPOSQVNNQOGTPV 552
 QY 94 SGRGFCVKMFYIKPIKYGPIKKEK 119
 DB 553 GS---VMVSYPTMSYOVPMTOGSG 574

RESULT 8
 F90571
 lipoprotein [imported] - Mycoplasma pulmonis (strain UAB CTIP)
 C:Species: Mycoplasma pulmonis
 C:Date: 24-May-2001 #sequence,revision 24-May-2001 #text_change 03-Aug-2001
 C:Accession: F90571

R:Chambaud, I.; Hellig, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Mos
 Nucleic Acids Res. 29, 2145-2153, 2001
 A:Title: The complete genome sequence of the murine respiratory pathogen Mycopla
 A:Reference number: A99512; MUID:21267165; PMID:11353084
 A:Accession: F90571
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-750 <KUR>
 A:Cross-references: GB:AL445566; PID:g14089892; PIDN:CAC13651.1; GSPDB:GN00153
 A:Experimental source: strain UAB CTIP
 C:Genetics:
 A:Gene: MYPU_4780
 A:Genetic code: SGC3

Query Match 9.2%; Score 74; DB 2; Length 750;
 Best Local Similarity 22.3%; Pred. No. 15;
 Matches 41; Conservative 25; Mismatches 56; Indels 62; Gaps 8;
 QY 12 LISTALGPMANICAGKSSNEIRTCDR--HGGGYSASORSQPHOGVDXCSAGSTYAP 69
 DB 456 LQSPYVSGWLD-----DRSIYSTKASDKTKDHLGEDILIEONKEVIAP 502
 QY 70 FTGMIVGQ---EKPYQ-----NNAINNGVRISSRGFCVKMF-- 103
 DB 503 FDGKIISAYAPSPYQAVGLGYITVLEWKKDLVGQIDQSVINDQLAETDRITYAFMHLN 562
 QY 104 -----YIKPIKYG-----PIKKEKGLTLPLQKVPYGIOSVHI-ENC 143
 DB 563 PSYLENGKLVQVSTAAIETTPATPKTKVKGVEIG-LVGEKKNNGMPPHVIIEVSLG 621
 QY 144 SSDE 147
 DB 622 STNP 625

RESULT 9
 AD0466
 Hypothetical protein YP03828 [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence,revision 02-Nov-2001 #text_change 02-Nov-2001
 C:Accession: AD0466
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tlball, R.W.; Holden, M.T.G.; Prenti
 deo-Parraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dou
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; B
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AD0466
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-350 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC93296.1; PID:g15981743; GSPDB:GN00175
 C:Genetics:
 A:Gene: YP03828

Query Match 9.1%; Score 73; DB 2; Length 350;
 Best Local Similarity 26.3%; Pred. No. 8.9;
 Matches 25; Conservative 17; Mismatches 39; Indels 14; Gaps 4;
 QY 52 PHOGVDXCSAGSTYAPF-----TGMTVGEKPYONKNAINGVRISSRGFCVKMFYIK 106
 DB 210 PH-----CHGEDLVASIKVTKDQGHITFEKLVGTNANISMTIIPFEKGYKLDIFPHE 263
 QY 107 PIKYGPIKKEKGLTLPLQKVPYGIOSVHIEN 141
 DB 264 PSRLR--VSSGDS-DTLIPESKONSFIIMTSIGIEN 295

RESULT 10
 XJHOFK
 formaldehyde transketolase (EC 2.2.1.3) - yeast (Pichia angusta)
 N:Alternate names: dihydroxyacetone synthase; glyceralone synthase
 C:Species: Pichia angusta

C>Date: 31-Mar-1993 #sequence_revision 30-Sep-1993 #text_change 20-Apr-2000
C:Accession: A23009; S30110
R:Janowicz, Z.A.; Eckart, M.R.; Drewke, C.; Roggenkamp, R.O.; Hollenberg, C.P.
Nucleic Acids Res. 13, 3043-3062, 1985
A>Title: Cloning and characterization of the DAS gene encoding the major methanol assim
A:Reference number: A23009; MUID:85215670
A:Accession: A23009
A:Molecule type: DNA
A:Residues: 1-693; 'RLPGREGKA' <JAN>
A:Cross-References: GB:X02424
A:Experimental source: ATCC 34438
A>Note: This sequence has been revised in reference S30110
R:Hansen, H.; Didion, T.; Thiemann, A.; Veenhuis, M.; Roggenkamp, R.
Mol. Gen. Genet. 235, 269-278, 1992
A>Title: Targeting sequences of the two major peroxisomal proteins in the methylotrophic
A:Reference number: S30110; MUID:93101130
A:Accession: S30110
A:Molecule type: DNA
A:Residues: 667-710 <HAN>
A:Cross-References: EMBL:X02424
Note: this is a revision to the sequence from reference A23009
Comment: This is the major methanol assimilatory enzyme from this methylotrophic organ

C:Genetics:
A:Gene: DAS
C:Superfamily: transketolase; thiamin pyrophosphate-binding domain homology
C:Keywords: peroxisome; thiamin pyrophosphate-binding domain homology #status atypical <TPB>
P:158-208/Domain: thiamin pyrophosphate-binding domain homology #status atypical
F:708-710/Region: peroxisome/glyoxysome location signal #status atypical

Query Match 9.1%; Score 73; DB 1; Length 710;
Best Local Similarity 24.1%; Pred. No. 18;
Matches 33; Conservative 21; Mismatches 49; Indels 34; Gaps 7;

QY 11 GLISTALAGWANTCAGKSSNEIRTCDRHGGCGYAGRSORPHQGVXLCASASTYVAPF 70
DB 525 GLFQKAVLEPFSSILS-LSRNEVL-----QYLASRAQRNRNAGYLLIEDAEMAEVQI 575
QY 71 TGMIVGOEKPYONKNAINGRISGRF-----CVKMFYIKPIKYGPI--KKGEKL 120
DB 576 IG--VGAEMEFADKAA-----KILGRKTRTRVLSIPCTRLFPDQSIGRISVLRKKDQRY 628
QY 121 GTLLPLOKXVYPGIOSHV 137
DB 629 PTYV-----VDGHV 637

RESULT 11
T36565
Probable penicillin-binding protein - Streptomyces coelicolor (fragment)

C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Nov-2000
C:Accession: T36565
R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21584
A:Accession: T36565
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-428 <SEP>
A:Cross-References: EMBL:AL079553; PIDN:CAB45563.1; GSPDB:GND0070; SCOEDB:SCH17.14
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCH17.14
C:Superfamily: penicillin-binding protein 1B

Query Match 9.0%; Score 72.5; DB 2; Length 428;
Best Local Similarity 38.9%; Pred. No. 12;
Matches 21; Conservative 6; Mismatches 22; Indels 5; Gaps 2;

QY 43 QYASQRSORPHQGVXLCASASTYVAPFTGMIV--GOEKPY--ONKNAINGV 91
DB 342 QESVHSLKDHVNRKSDSVAAASTIVEPTGRIVAMGQSKPYGKNETETINTSY 395

RESULT 12
AC2451
hypothetical protein all5163 [imported] - Anabaena sp. (strain PCC 7120)

C:Species: Anabaena sp.
A>Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AC2451
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Ir
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tab
DNA Res. 8, 205-213, 2001
A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacteri
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AC2451
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <NUR>
A:Cross-References: GB:BA000019; PIDN:BA076862.1; PID:g17134301; GSPDB:GND0179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all5163

Query Match 8.9%; Score 71.5; DB 2; Length 295;
Best Local Similarity 29.6%; Pred. No. 11;
Matches 34; Conservative 14; Mismatches 40; Indels 27; Gaps 7;

QY 53 HOGVDXCSAGSYV--YAPFTGMIVGOEKPYONKNAIN-NGVIRISGRFCVKMFYIKPIK 109
DB 190 HSGVDLLAAGTPEALAPGTVFAKREGTYGNLYINNG-----GLQSRVAHDSIN 243
QY 110 YK--GPIKRGKGLGTL-----LPLOKVPYGIOSHVHIENCSS-----DPTAYL 151
DB 244 VKVGKQVNCQQLLTGTGTGTPAK-----QPHLHFVYVSSSLGVAEMPKDYL 293

RESULT 13
T01351
subtilisin-like proteinase homolog F6N15.3 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 20-Sep-1999
C:Accession: T01351
R:Ryan, E.; Edwards, J.; Pape, K.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of A. thaliana F6N15.
A:Reference number: Z14297
A:Accession: T01351
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-706 <RYA>
A:Cross-References: EMBL:AF069299; NID:g3193311; PID:g3193320
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:Intons: 10/2; 43/1; 91/1; 126/2; 297/3; 330/2; 418/2; 454/3; 526/1
A>Note: F6N15.3
C:Superfamily: subtilisin-like proteinase ag12; subtilisin homology

Query Match 8.9%; Score 71.5; DB 2; Length 706;
Best Local Similarity 25.0%; Pred. No. 26;
Matches 33; Conservative 21; Mismatches 35; Indels 43; Gaps 8;

QY 8 LLAG-----LISTALAGPWANICAGKSSNEIRTCDRHGGCGYSAORSORPHQGV 57
DB 206 ILAGFEAIIHDGVETIISISIGPIADY-----SSDSI-----SVGSFHAMR-----KGIL 250
QY 58 XLCASAGSYVAPFTGMIV-----GOEKPYONKNAINGVIRISGRFCVKMFYI 105
DB 251 TVASAGND--GPSGTVNHEPWILTVAASGIDRTFKSIIDGNGSFSQMG--ISMFSP 306

QY 106 KPIKYGKPIKKG 117
DB 307 KAKSY--PLVSG 316

RESULT 14

E97836

probable periplasmic protein [imported] - Rickettsia conorii (strain Malish 7)

C:Species: Rickettsia conorii

C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001

C/Accession: E97836

R/Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R

Science 293, 2093-2098, 2001

A>Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A/Reference number: A97700; M01D:21442074; PMID:1157893

A/Accession: E97836

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-284 <KUR>

A/Cross-references: GB:AE006914; PIDN:AAL03631.1; PID:915620216; GSPDB:GN00173

C/Genetics:

A/Gene: RCI093

Query Match 8.8%; Score 71; DB 2; Length 284;

Best Local Similarity 24.0%; Pred. No. 12;

Matches 24; Conservative 16; Mismatches 32; Indels 28; Gaps 5;

OY 53 HQGVDXLCSAGSTVYAPFTGMIV-GOEKPYQNKNAINGVRISGRFCVVMFYTKPIKY 110
 DB 181 HSGIDLQAKKAPIYAAAGIVIKARAPDYGNFVEIKH-----GRKFVTYVAHLKEMSV 235

OY 111 K-GPIKKGKIKGTLLPLQKVPGIQ-----SHVHIE 140
 DB 236 KEQNKIKRQ-----FIGIGSTGNATGEHLHFE 264

RESULT 15

T05614

hypothetical protein F9D16.290 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 17-Mar-2000

C/Accession: T05614

R/Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hohnsels, J.; Mewes, H.W.; Meyer, K.F.

submitted to the Protein Sequence Database, February 1999

A/Reference number: 215419

A/Accession: T05614

A/Molecule type: DNA

A/Residues: 1-444 <BEV>

A/Cross-references: EMBL:AL035394

A/Experimental source: cultivar Columbia; BAC clone F9D16

C/Genetics:

A/Map position: 4

A/Introns: 110/3; 151/1; 209/2; 233/1

A/Note: F9D16.290

C/Superfamily: polygalacturonase

Query Match 8.7%; Score 70; DB 2; Length 444;

Best Local Similarity 25.7%; Pred. No. 23;

Matches 38; Conservative 14; Mismatches 54; Indels 42; Gaps 8;

OY 23 NICAGKSSNEIRTCQ-----RHCGOYSQSRPHQGVDXLCSAGSTVYAPFTGMIVGQ 77
 DB 237 NVCT--EDSYISTGDDLVAIKSGMDYGIAYG-RPSSNITIRRTGS--SPFAGIATGS 290

OY 78 EKPYQNNAINNGVRISGRFCVKM-----FYIKPIK----YKGPICKKGKIG----- 121
 DB 291 ETSGGIKNIITAEHITLSNMGVGVNITKNGRGYIKNIKISDVYVDPAKYGIKTAGDTGD 350

OY 122 -----TLPLQKVYPGIQSHVHIE 141
 DB 351 HPDENINPNALPVVK-----GIHKN 371

Search completed: October 21, 2002, 16:34:18
 Job time : 19 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2002, 16:26:52 ; Search time 12 Seconds

(without alignments)
487.221 Million cell updates/sec

Title: US-09-868-953-1

Perfect score: 804

Sequence: 1 MFSTKALLAGLSTALACP.....GIQSHVHENCDSPTAYL 151

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	802	99.8	151	1	LCT2_HUMAN
2	698	86.8	151	1	LCT2_BOVIN
3	681	84.7	151	1	LCT2_MOUSE
4	429	53.4	326	1	MIMI_CHICK
5	76.5	9.5	475	1	YEB4_HAERIN
6	73	9.1	710	1	DAS_PICAN
7	71	8.8	779	1	TRF3_THEVO
8	70.5	8.8	680	1	Z334_HUMAN
9	69	8.6	1330	1	XDH_RAT
10	68	8.5	184	1	YT63_CAEEL
11	68	8.5	762	1	MUS2_AQUAE
12	67.5	8.4	355	1	COLL_ARATH
13	67.5	8.4	535	1	VG50_HSVSA
14	67.5	8.4	625	1	TR11_MOUSE
15	67	8.3	398	1	MUB1_XENLA
16	65	8.1	271	1	YLBF_ECOLI
17	65	8.1	312	1	PGLR_AGRPU
18	65	8.1	317	1	TBP2_SHEEP
19	65	8.1	1388	1	RPOD_TOBAC
20	64.5	8.0	474	1	FAY_RABIT
21	64.5	8.0	541	1	SECD_MYCLE
22	64.5	8.0	2241	1	TEGU_HCMVA
23	64	8.0	593	1	MTE2_HUMAN
24	64	8.0	746	1	GYP7_YEAST
25	64	8.0	1477	1	KELC_DROME
26	63.5	7.9	551	1	LEM2_RABIT
27	63	7.8	417	1	KCRD_HUMAN
28	63	7.8	490	1	HDVD_CLOAM
29	63	7.8	964	1	PTSA_STRMU
30	63	7.8	947	1	PM19_CHLPN
31	62.5	7.8	466	1	FA7_HUMAN
32	62.5	7.8	691	1	OAT6_HUMAN
33	62	7.7	109	1	CYC_TETPY

34	62	7.7	221	1	GPII_ZINOF	P82474 zingiber of
35	62	7.7	274	1	VPRT_MMTVB	P10271 mouse mamma
36	62	7.7	499	1	DHAS_CHICK	O93344 gallus gall
37	62	7.7	775	1	YTX1_XENLA	P14380 xenopus lae
38	62	7.7	948	1	HMDH_SCHNA	P16237 schistosoma
39	62	7.7	1196	1	BXCN_CLOBO	P46081 clostridium
40	62	7.7	2109	1	RRPL_VSVTH	P13615 vesicular s
41	62	7.7	3412	1	POLG_TBETS	P07720 t genome po
42	61.5	7.6	362	1	FTBP_ADDE8	P36845 human adeno
43	61.5	7.6	391	1	HISZ_BACSU	O34459 bacillus su
44	61.5	7.6	391	1	SSRI_MOUSE	P30873 mus musculu
45	61.5	7.6	493	1	ISTP_STASI	P10547 staphylococ

ALIGNMENTS

RESULT 1	ID	SEQUENCE	STANDARD	PRT	151 AA.
LCT2_HUMAN	AC	O14960; O14565;			
LCT2_HUMAN	DT	15-JUL-1999 (Rel. 38, Created)			
LCT2_HUMAN	DT	15-JUL-1999 (Rel. 38, Last sequence update)			
LCT2_HUMAN	DT	16-OCT-2001 (Rel. 40, Last annotation update)			
LCT2_HUMAN	DE	Leukocyte cell-derived chemotaxin 2 precursor (hLECT2).			
LCT2_HUMAN	GN	LCT2.			
LCT2_HUMAN	OS	Homo sapiens (human).			
LCT2_HUMAN	OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
LCT2_HUMAN	OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
LCT2_HUMAN	NCBI	TaxID=9606;			
LCT2_HUMAN	RP	SEQUENCE FROM N.A. AND VARIANT ILE-58.			
LCT2_HUMAN	RC	TISSUE=Liver;			
LCT2_HUMAN	RC	MEDLINE=98193133; PubMed=9524238;			
LCT2_HUMAN	RA	Yamagoe S., Mizuno S., Suzuki K.;			
LCT2_HUMAN	RT	"Molecular cloning of human and bovine LECT2 having a neutrophil			
LCT2_HUMAN	RT	chemotactic activity and its specific expression in the liver.";			
LCT2_HUMAN	RL	Biochim. Biophys. Acta 1396:105-113(1998).			
LCT2_HUMAN	RP	SEQUENCE FROM N.A.			
LCT2_HUMAN	RC	TISSUE=Periphereal blood;			
LCT2_HUMAN	RC	MEDLINE=98207247; PubMed=9545637;			
LCT2_HUMAN	RA	Yamagoe S., Kameoka Y., Hashimoto K., Mizuno S., Suzuki K.;			
LCT2_HUMAN	RT	"Molecular cloning, structural characterization, and chromosomal			
LCT2_HUMAN	RT	mapping of the human LECT2 gene.";			
LCT2_HUMAN	RL	Genomics 48:324-329(1998).			
LCT2_HUMAN	RP	SEQUENCE FROM N.A.			
LCT2_HUMAN	RC	MEDLINE=99048897; PubMed=9832057;			
LCT2_HUMAN	RA	Nagai H., Hamada T., Uchida T., Yamagoe S., Suzuki K.;			
LCT2_HUMAN	RT	"Systemic expression of a newly recognized protein, LECT2, in the			
LCT2_HUMAN	RT	human body.";			
LCT2_HUMAN	RL	Pathol. Int. 48:882-886(1998).			
LCT2_HUMAN	RP	SEQUENCE FROM N.A.			
LCT2_HUMAN	RC	Kimmerly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,			
LCT2_HUMAN	RA	Kadher K., Miguel T., Miller C., Pitluck S., Pollard M., Rojkeski H.,			
LCT2_HUMAN	RT	Submitted (May-1998) to the EMBL/GenBank/DBJ databases.			
LCT2_HUMAN	RL	Submitted (May-1998) to the EMBL/GenBank/DBJ databases.			
LCT2_HUMAN	RP	SEQUENCE FROM N.A.			
LCT2_HUMAN	RC	Strong C., Blewald T., Tin-Wollam A., Duckels G.;			
LCT2_HUMAN	RA	Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.			
LCT2_HUMAN	RP	SUBCELLULAR LOCATION.			
LCT2_HUMAN	RA	MEDLINE=97031488; PubMed=8877413;			
LCT2_HUMAN	RT	Yamagoe S., Yamakawa Y., Matsuo Y., Minowada J., Mizuno S., Suzuki K.;			
LCT2_HUMAN	RT	"Purification and primary amino acid sequence of a novel neutrophil			
LCT2_HUMAN	RT	chemotactic factor LECT2.";			
LCT2_HUMAN	RL	Immunol. Lett. 52:9-13(1996).			
LCT2_HUMAN	CC	-I- FUNCTION: Has a neutrophil chemotactic activity. Also a positive			
LCT2_HUMAN	CC	regulator of chondrocyte proliferation.			
LCT2_HUMAN	CC	-I- SUBCELLULAR LOCATION: Cytoplasmic and secreted.			

CC -1- TISSUE SPECIFICITY: Highly expressed in adult and fetal liver and
CC weakly in testis. Not expressed in bone marrow.
CC -1- INDUCTION: By phytohemagglutinin (PHA).
CC -1- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D63521; BAA23609.1; -
DR EMBL: AB007546; BAA25669.1; -
DR EMBL: AC004763; AAC17734.1; -
DR EMBL: AC004248; AAB66905.1; -
DR MIM: 602882; -
KW Chemotaxis; Signal; Polymorphism.
FT SIGNAL 1 18 BY SIMILARITY.
FT CHAIN 19 151 LEUCOCYTE CELL-DERIVED CHEMOTAXIN 2.
FT VARIANT 58 58 V->I
SQ SEQUENCE 151 AA; 16376 MW; A4267F0A24E69631 CRC64;
Query Match 99.8%; Score 802; DB 1; Length 151;
Best Local Similarity 99.3%; Pred. No. 1e-77;
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MESTKALLAGLISTALAGPANTCAGKSSNEIRCTDRHGGCGOYSAORSORPHQGVXLC 60
DB 1 MESTKALLAGLISTALAGPANTCAGKSSNEIRCTDRHGGCGOYSAORSORPHQGVXLC 60
QY 61 SAGSTYVAPPTGMIIVGGEKPYONKNAIINGVRIISGRCVKNFYIKPIYKGIKKEKL 120
DB 61 SAGSTYVAPPTGMIIVGGEKPYONKNAIINGVRIISGRCVKNFYIKPIYKGIKKEKL 120
QY 121 GTLLPLQKVPYGIQSHVHIENCSSDPYVYL 151
DB 121 GTLLPLQKVPYGIQSHVHIENCSSDPYVYL 151
RESULT 2
LCT2_BOVIN STANDARD: PRT; 151 AA.
ID LCT2_BOVIN
AC O62644;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Leukocyte cell-derived chemotaxin 2 precursor (bLECT2) (Chondromodulin
DE II) (bCHM-II).
GN LECT2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98193133; PubMed=9524238;
RA Yamagoe S., Mizuno S., Suzuki K.;
RT "Molecular cloning of human and bovine LECT2 having a neutrophil
RT chemotactic activity and its specific expression in the liver";
RL Biochim. Biophys. Acta 1396:105-113(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=99160594; PubMed=10050029;
RA Shukunami C., Kondo J., Wakai H., Takahashi K., Inoue H., Kamizono A.,
RA Hiraki Y.;
RT "Molecular cloning of mouse and bovine chondromodulin-II cDNAs and the
RT growth-promoting actions of bovine recombinant protein.";

RL J. Biochem. 125:436-442(1999).
RN [3]
RP SEQUENCE OF 19-151.
RC TISSUE=Epiphyseal cartilage;
RX MEDLINE=96594331; PubMed=8796437;
RA Hiraki Y., Inoue H., Kondo J., Kamizono A., Yoshitake Y.,
RA Shukunami C., Suzuki F.;
RT "A novel growth-promoting factor derived from fetal bovine cartilage,
RT chondromodulin II. Purification and amino acid sequence.";
RL J. Biol. Chem. 271:22657-22662(1996).
CC -1- FUNCTION: Has a neutrophil chemotactic activity. Also a positive
CC regulator of chondrocyte proliferation.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB001350; BAA25302.1; -
DR EMBL: D89011; BAB18616.1; -
KW Chemotaxis; Signal.
FT SIGNAL 1 18
FT CHAIN 19 151 LEUCOCYTE CELL-DERIVED CHEMOTAXIN 2.
SQ SEQUENCE 151 AA; 16320 MW; 2A1DE8F5B28A3D8 CRC64;
Query Match 86.8%; Score 698; DB 1; Length 151;
Best Local Similarity 86.1%; Pred. No. 9.9e-67;
Matches 130; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
QY 1 MESTKALLAGLISTALAGPANTCAGKSSNEIRCTDRHGGCGOYSAORSORPHQGVXLC 60
DB 1 MESTGTLIALALISPAAGPAITCAGKSSNEIRCTDRHGGCGOYSAORSORPHQGVXLC 60
QY 61 SAGSTYVAPPTGMIIVGGEKPYONKNAIINGVRIISGRCVKNFYIKPIYKGIKKEKL 120
DB 61 SAGSTYVAPPTGMIIVGGEKPYONKNAIINGVRIISGRCVKNFYIKPIYKGIKKEKL 120
QY 121 GTLLPLQKVPYGIQSHVHIENCSSDPYVYL 151
DB 121 GTLLPLQKVPYGIQSHVHIENCSSDPYVYL 151
RESULT 3
LCT2_MOUSE STANDARD: PRT; 151 AA.
ID LCT2_MOUSE
AC O88803; O88804; O90WN3; O92337;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Leukocyte cell-derived chemotaxin 2 precursor (Chondromodulin II)
DE (CHM-II).
GN LECT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANT VAL-129.
RC STRAIN=BALB/C; TISSUE=Liver;
RX MEDLINE=98382586; PubMed=9714793;
RA Yamagoe S., Watanabe T., Mizuno S., Suzuki K.;
RT "The mouse LECT2 gene: cloning of cDNA and genomic DNA, structural
RT characterization and chromosomal localization.";
RL Gene 216:171-178(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=Swiss Webster / NIH Swiss; TISSUE=Embryo, and Liver;
RX MEDLINE=99160594; PubMed=10050029;

RA Shukunami C., Kondo J., Wakai H., Takahashi K., Inoue H., Kamizono A.,
 RA Hiraki Y.;
 RT "Molecular cloning of mouse and bovine chondromodulin-II cDNAs and the
 RT growth-promoting actions of bovine recombinant protein.";
 RL J. Biochem. 125:436-442(1999).
 CC
 CC -1- FUNCTION: Has a neutrophil chemotactic activity. Also a positive
 CC regulator of chondrocyte proliferation.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1/LECT2 (shown here) and
 CC 2/LECT20: are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Highly expressed in liver and weakly in
 CC testis. Not expressed in heart, brain, spleen, lung, skeletal
 CC muscle and kidney.
 CC -1- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AB009687; BA33383.1; -
 CC DR EMBL: AB009688; BA33384.1; -
 CC DR EMBL: AB009689; BA33385.1; -
 CC DR EMBL: AB009689; BA33386.1; -
 CC DR EMBL: AF035161; AF13302.1; -
 CC DR MGD: MG1:1278342; LECT2.
 CC
 CC Chemotaxis; Signal; Alternative splicing.
 CC SIGNAL 1 18 BY SIMILARITY.
 CC CHAIN 19 151 LEUCOCYTE CELL-DERIVED CHEMOKINE 2.
 CC FT VARSPLIC 98 151 FCVRIFYIKPIKYGSIKKGKIGTLPLPKIYGIOSHVH
 CC FT VENCSDSDPTAYL -> ORLOAHATTILNPTCYWDKIQIPR
 CC FT PTFELQCNFLH (IN ISOFORM LECT20).
 CC
 CC VARIANT 129 129 I -> V
 CC FT SEQUENCE 151 AA; 16405 MW; 18AF444046B7AE8E CRC64;
 CC SQ

Query Match 84.7%; Score 681; DB 1; Length 151;
 Best Local Similarity 82.1%; Pred. No. 6.2e-65;
 Matches 124; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

QY 1 MESTKALLAGLSTALAGPWANICAGKSSNEIFTCDRHCGGYSAGRSORPHOGVDXLC 60
 DB 1 MPTTILISALLSSALAGPWANICAGKSSNEIFTCDSYCGGYSAQORPHRNGVDVLC 60
 QY 61 SAGSYVAPFTGMIVGOEKPYONKNAINGVRISGRGFCVKMFYIKPIKYGKPKKEKL 120
 DB 61 SDGSVYVAPFTGKIVGOEKPRRNKNAIDGIRLSGRGFCVKIPIKIKYKSIKKEKL 120
 QY 121 GTLLPLQKYPYGIOSHVIHENCSDSDPTAYL 151
 DB 121 GTLLPLQKYPYGIOSHVIHENCSDSDPTAYL 151

RESULT 4
 MIM1_CHICK STANDARD; PRT; 326 AA.
 ID MIM1_CHICK
 AC P08940;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Myeloid protein-1 precursor (P33).
 GN MIM-1.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 CC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90090611; Pubmed=2688896;
 RA Ness S.A., Marknell A., Graf T.;

RT "The v-myb oncogene product binds to and activates the promyelocyte-
 RT specific mim-1 gene."
 RL Cell 59:1115-1125(1989).
 RN [2]
 RP PARTIAL SEQUENCE; AND SEQUENCE OF 295-299 FROM N.A.
 RC STRAIN-WHITE LEHORN;
 RA MEDLINE=93011981; Pubmed=1397316;
 RX Yamada K., Tsuchiya M., Mishima K., Shimoyama M.;
 RT "P33, an endogenous target protein for arginine-specific ADP-
 RT ribosyltransferase in chicken polymorphonuclear leukocytes, is highly
 RL homologous to mim-1 protein (myb-induced myeloid protein-1).";
 RL FEBS Lett. 311:203-205(1992).
 CC -1- FUNCTION: NOT KNOWN.
 CC -1- SUBCELLULAR LOCATION: GRANULES OF PROMYELOCYTES.
 CC -1- INDUCTION: BY THE MYB ONCOGENE.
 CC -1- PFM: SUBSTRATE FOR ARGININE-SPECIFIC ADP-RIBOSYLTRANSFERASE.
 CC -1- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M29448; AAA48954.1; -
 CC DR EMBL: M29449; AAA48958.1; -
 CC DR PIR: A33755; A33755.
 CC DR PIR: S29229; S29229.
 CC
 CC Repeat; Signal.
 CC SIGNAL 1 23 PROBABLE.
 CC CHAIN 24 326 MYELOID PROTEIN-1.
 CC FT REPEAT 28 162
 CC FT REPEAT 177 312
 CC FT REPEAT 297 297
 CC FT CONFLICT 297 297 I -> Y (IN REF. 1).
 CC FT SEQUENCE 326 AA; 35636 MW; B8C4742EF0D3BAD3 CRC64;
 CC SQ

Query Match 53.4%; Score 429; DB 1; Length 326;
 Best Local Similarity 50.3%; Pred. No. 5.8e-38;
 Matches 80; Conservative 26; Mismatches 39; Indels 14; Gaps 2;

QY 7 LLLAGLSTALAGP-----MANICAGKSSNEIFTCDRHCGGYSAGRSOR--P 52
 DB 7 IALLSVSTAFARQMEVHPQQQGRHMQDLCGPNRRIRGCDRIGGNGAGSAGQGEK 66
 QY 53 HGVYDLCSAGSYVAPFTGMIVGOEKPYONKNAINGVRISGRGFCVKMFYIKPIKYGK 112
 DB 67 HKGVYDLCTDGSIVYAPFSGQLSGPIRPFHNNNAIDDGVOISGSGCVKILCIPRYHG 126
 QY 113 PIKKEKLTGTLPLQKYPYGIOSHVIHENCSDSDPTAYL 151
 DB 127 QIQKGQDLGRMLPMQKVPFGIVSHIHVENCSDSDPTAYL 165

RESULT 5
 YEBA_HAELIN STANDARD; PRT; 475 AA.
 ID YEBA_HAELIN
 AC P44693;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein H10409.
 GN H10409.
 OS Haemophilus influenzae.
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.
 CC NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; Pubmed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,

RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-T., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Usterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Georgagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus
 RT influenzae Rd.,"
 RL Science 269:496-512(1995).
 CC -1- FUNCTION: COULD BE INVOLVED IN CELL WALL DEGRADATION OR
 CC FORMATION (BY SIMILARITY).
 CC -1- SIMILARITY: STRONG, TO E.COLI YEBA.
 CC -1- SIMILARITY: TO STAPHYLOCOCCUS LYSOSTAPHIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U32724; AAC22068.1; -.
 DR TIGR: H10409; -.
 DR InterPro: IPR002886; Peptidase_M37.
 DR Pfam: PF01551; Peptidase_M37.1.
 DR Pfam: PF01551; Peptidase_M37.1.
 DR Hypothetical protein; Cell wall; 143C10F9223939D CRC64;
 KW SEQUENCE 475 AA; 53255 MW; 143C10F9223939D CRC64;
 SO

Query Match 9.5%; Score 76.5; DB 1; Length 475;
 Best Local Similarity 28.7%; Pred. No. 1.5;
 Matches 27; Conservative 15; Mismatches 41; Indels 11; Gaps 4;

QY 51 RPHQGVXICLSAGSTVYAPTGMIVGQEPYONKNAINGVRISGRGFCVKMYKP-I 108
 DB 346 RPHQGVXICLSAGSTVYAPTGMIVGQEPYONKNAINGVRISGRGFCVKMYKP-I 108
 QY 109 KYKGPTRKKGKGLTLPLQKYPYQIQS--HVHIE 140
 DB 404 KAGQTVKKGKRIAL-----SGNTGISTGPHLHYE 432

RESULT 6

DAS_PICAN STANDARD; PRT; 710 AA.

AC P06834;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Dihydroxy-acetone synthase (EC 2.2.1.3) (DHAS) (Formaldehyde
 DE transketolase) (Glycerone synthase).
 GN DAS.
 OS Pichia angusta (Yeast) (Hansenula polymorpha).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4905;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 34438;
 RX MEDLINE=85215670; PubMed=2987872;
 RA Janowicz Z.A., Eckart M.R., Drewke C., Roggenkamp R.O.,
 RA Hollenberg C.P., Matz J., Ledebert A.M., Visser C.T.;
 RT "Cloning and characterization of the DAS gene encoding the major
 RT methanol assimilatory enzyme from the methylotrophic yeast Hansenula
 RT polymorpha,"
 RL Nucleic Acids Res. 13:3043-3062(1985).
 [2]
 RP REVISIONS, SEQUENCE OF 667-710 FROM N.A.
 RC STRAIN-ATCC 34438;
 RT MEDLINE=93101130; PubMed=1465101;
 RA Hansen H., Didion T., Thiemann A., Veenhuis M., Roggenkamp R.O.;

RT "Targeting sequences of the two major peroxisomal proteins in the
 RT methylotrophic yeast Hansenula polymorpha,"
 RL Mol. Gen. Genet. 235:269-278(1992).
 CC -1- FUNCTION: THIS IS THE MAJOR METHANOL ASSIMILATORY ENZYME FROM
 CC THE METHYLOTROPHIC HANSENULA POLYMORPHA.
 CC -1- CATALYTIC ACTIVITY: D-xylulose 5-phosphate + formaldehyde =
 CC glyceraldehyde 3-phosphate + glyceroine.
 CC -1- COFACTOR: THIAMINE PYROPHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: Peroxisomal.
 CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X02424; CAA36276.1; -.
 DR PIR: A23009; XJHOK.
 DR HSSP: P23254; TTRK.
 DR InterPro: IPR000360; Transketolase.
 DR Pfam: PF00456; transketolase; 1.
 DR Pfam: PF02780; transketolase; 1.
 DR Pfam: PF02779; transketolase; 1.
 DR PROSITE: PS00801; TRANSKETOLASE_1; 1.
 DR PROSITE: PS00802; TRANSKETOLASE_2; 1.
 DR PROSITE: PS00342; MICROBODIES_CTER; 1.
 DR Methanol utilization; Peroxisome; Transferrase; Thiamine pyrophosphate.
 KW SITE 708
 FT SEQUENCE 710 AA; 78842 MW; E4424B6D3CC8B0B CRC64;
 SO

Query Match 9.1%; Score 73; DB 1; Length 710;
 Best Local Similarity 24.1%; Pred. No. 5.3;
 Matches 33; Conservative 21; Mismatches 49; Indels 34; Gaps 7;

QY 11 GLISTALAGWANICAKSSNEIFTCDRHGGCGQYASQSRPHQGVXICLSAGSTVYAPF 70
 DB 525 GLFQKAVELPFSILTS--LSRNEVL-----QYLASRAQRRRNAGYILEDENAEVQI 575
 QY 71 TGMIVGQEPYONKNAINGVRISGRG-----CYKMPYIKPKYKGP--KKGKGL 120
 DB 576 IG--VGAEMEFADKAA-----KILGRKRTRYLSIPCTRLPDEQSIGRYSVLRKGRQY 628
 QY 121 GTLLPLQKYPYQIQSHV 137
 DB 629 PTYV-----VDGHV 637

RESULT 7

TRF3_THEVO STANDARD; PRT; 779 AA.

AC Q97A16;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Tricoin protease interacting factor F3 (EC 3.4.11.-).
 GN TRF3 OR TVG0817891.
 OS Thermoplasma volcanium.
 OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmataceae;
 OC Thermoplasma.
 OX NCBI_TaxID=50339;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GSSI / DSM 4299 / JCM 9571;
 RX MEDLINE=20570466; PubMed=11121031;
 RA Kawashima T., Ameno N., Koike H., Makino S.-I., Higuchi S.,
 RA Kawashima-Ohta Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
 RA Nunoshima T., Yamamoto Y., Aizawa H., Makino K., Suzuki M.;
 RT "Archaeal adaptation to higher temperatures revealed by genomic
 RT sequence of Thermoplasma volcanium,"
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).

CC -1- FUNCTION: Proteases F1, F2 and F3 degrade oligopeptides produced
 CC by tricorn (themselves probably produced by the proteasome),
 CC yielding free amino acids (By similarity).
 CC -1- COFACTOR: Binds one zinc ion (By similarity).
 CC -1- SUBUNIT: Part of the tricorn proteolytic complex (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AP000993; BAB59956.1; -
 CC DR MEROPS: M01.021; -
 DR InterPro: IPR001930; Aladiptase.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01433; Peptidase_M1; 1.
 DR PRINTS: PR00756; ALADIPTASE.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metallopeptase; Aminopeptidase; zinc; Complete proteome.
 FT METAL SITE 266 266 ZINC (CATALYTIC) (By similarity).
 FT ACT SITE 267 267 ZINC (CATALYTIC) (By similarity).
 FT METAL 270 270 ZINC (CATALYTIC) (By similarity).
 FT METAL 289 289 ZINC (CATALYTIC) (By similarity).
 FT ACT SITE 352 352 PROTON DONOR (POTENTIAL).
 SQ SEQUENCE 779 AA; 88068 MW; 68C227F2ECCECB201 CRC64;
 Query Match 8.88; Score 71; DB 1; Length 779;
 Best Local Similarity 26.48; Pred. No. 9.5;
 Matches 33; Conservative 21; Mismatches 59; Indels 12; Gaps 4;
 QY 24 ICAGKSSNEIRTCDRHCGCGOYSAORSQPHGVNDLCSAGSYVYAFPTGIVGQERYON 83
 DB 590 IAGKASRLALDEAVC-ETLAPRFNSFEDQTPPELKSATAYALSTGDKMVERKRS 648
 QY 84 KNAINGVR-ISGRG---FCVKNFVYKPIKIKGPIKIGKELG-----TLPLQKYYPG 132
 DB 649 LDRDEKVKIKISGFGKLSSTDLVSVMIEGKIKKQDMLSFYLSALEMAQREVIYSN 708
 QY 133 IOSHV 137
 DB 709 LENVV 713
 RESULT 8
 ID 334_HUMAN STANDARD; PRT; 680 AA.
 AC O9HCZ1; O9NVW4;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Zinc finger protein 334.
 GN ZNF334.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosich K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knighs A., Laird G.K., Lawlor S.,
 RA Lehtsalaho M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., Mclay K., McMurray A.A.,
 RA Milne S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Philimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming J., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:665-871(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiraori A., Sudo H.,
 RA Matsunuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishi S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masubo Y.,
 RA Ninomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project";
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC
 CC -1- SIMILARITY: CONTAINS 1 KRAB DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AL034424; CAC18896.1; -
 CC EMBL: AL034475; CAC03544.1; -
 CC EMBL: AK001331; BAA91630.1; -
 CC HSSP: P03001; 1TF6.
 DR InterPro: IPR001909; KRAB.
 DR InterPro: IPR000822; ZnF-C2H2.
 DR Pfam: PF00096; zf-C2H2; 14.
 DR PRINTS: PR00048; ZINCINGER.
 DR SMART: SM00349; KRAB; 1.
 DR SMART: SM00355; ZNF_C2H2; 14.
 DR PROSITE: PS00805; KRAB; 1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 14.
 DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 14.
 KW Transcription regulation; zinc-finger; Metal-binding; Nuclear protein;
 KW DNA-binding; Repeat.
 FT DOMAIN 10 81 KRAB.
 FT ZINC FINGERS.
 FT ZN_FING 237 678
 FT ZN_FING 237 259
 FT ZN_FING 265 287
 FT ZN_FING 293 315
 FT ZN_FING 321 343
 FT ZN_FING 349 371
 FT ZN_FING 377 399
 FT ZN_FING 405 427
 FT ZN_FING 433 455
 FT ZN_FING 461 483
 FT ZN_FING 483 505
 FT ZN_FING 511 533
 FT ZN_FING 544 566
 FT ZN_FING 572 594
 FT ZN_FING 600 622
 FT ZN_FING 628 650
 FT ZN_FING 656 678

FT CONFLICT 118 118 K -> E (IN REF. 2).
 FT CONFLICT 426 426 S -> G (IN REF. 2).
 SQ SEQUENCE 680 AA; 79648 MW; 77DB5086B9261DC6 CRC64;
 Query Match
 Best Local Similarity 21.4%; Pred. No. 9.3;
 Matches 28; Conservative 15; Mismatches 45; Indels 43; Gaps 4;
 8.8%; Score 70.5; DB 1; Length 680;
 OY 27 GKSNEIRTCDRHGGCGQYSNORSQRPQGVDXLCSAGSTVYAPFTGMTVGGEKPYQNKNA 86
 DB 566 GGRPYECNECGKTECFCKSFVEHQRTHTG-----EKPYECNEC 605
 OY 87 INNGVVISGRGFCVKMY-----IKPIKYGPPIKKGKGLTLLPLQKYVPGDLSHV 137
 DB 606 -----GKSFCHKSAFVRHRIHTGKEPYECNOCGYRRLMTLTHOKIHTGKPY- 656
 OY 138 HIENCSDSPT 148
 DB 657 ---ECKKCEKT 664
 RESULT 9
 XDH_RAT STANDARD; PRT; 1330 AA.
 ID XDH_RAT 063157;
 AC P22985;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Xanthine dehydrogenase/oxidase [includes: Xanthine dehydrogenase (EC 1.1.1.204) (XD); Xanthine oxidase (EC 1.1.3.22) (XO) (Xanthine oxidoreductase)].
 GN XDH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=90354396; PubMed=2387845;
 RA Amaya Y., Yamazaki K.-I., Sato M., Noda K., Nishino T., Nishino T.;
 RT NAD-dependent conversion of xanthine dehydrogenase from the
 RT NAD-dependent type to the O2-dependent type. Amino acid sequence of
 RT rat liver xanthine dehydrogenase and identification of the cleavage
 RT sites of the enzyme protein during irreversible conversion by
 RT trypsin.";
 RL J. Biol. Chem. 265:14170-14175(1990).
 RN [2]
 RP SEQUENCE OF 1-54 FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=94268906; PubMed=8208609;
 RA Chow C.W., Clark M., Rinaldo J., Chalkley R.;
 RT Identification of the rat xanthine dehydrogenase/oxidase promoter.";
 RL Nucleic Acids Res. 22:1846-1854(1994).
 CC -1- FUNCTION: THIS ENZYME CAN BE CONVERTED FROM THE DEHYDROGENASE FORM
 CC (D) TO THE OXIDASE FORM (O) IRREVERSIBLY BY PROTEOLYSIS OR
 CC REVERSIBLY THROUGH THE OXIDATION OF SULFHYDRYL GROUPS.
 CC -1- CATALYTIC ACTIVITY: Xanthine + NAD(+) + H(2)O = urate + NADH.
 CC -1- CATALYTIC ACTIVITY: Xanthine + H(2)O + O(2) = urate + H(2)O(2).
 CC -1- COFACTOR: FAD, MOLYBDOTERIN, AND TWO 2FE-2S CLUSTERS.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Peroxisomal.
 CC -1- INDUCTION: BY INTERFERON.
 CC -1- SIMILARITY: TO OTHER XANTHINE DEHYDROGENASES, AND LIMITED TO
 CC OTHER EUKARYOTIC MOLYBDOTERIN ENZYMES SUCH AS NITRATE REDUCTASE
 CC AND SULFITE OXIDASE.
 CC -1- SIMILARITY: TO 2FE-2S FERREDOXINS IN THE N-TERMINAL DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J05579; AAA42349.1; -;
 DR EMBL: U08122; AAA18869.1; -;
 DR EMBL: U08120; AAA18869.1; JOINED.
 DR EMBL: U08121; AAA18869.1; JOINED.
 DR PIR: A37810; A37810.
 DR HSSP: Q46509; IALO.
 DR InterPro: IPR002888; 2Fe-2S_BD.
 DR InterPro: IPR000564; 2Fe2S_ferredoxin.
 DR InterPro: IPR000674; Aldxan_dh_C.
 DR InterPro: IPR000572; Euk oxidored_mol_yb.
 DR InterPro: IPR010041; ferredoxin.
 DR InterPro: IPR002346; dehydrog_mol_yb.
 DR Pfam: PF01315; Ald_Xan_dh_C; 1.
 DR Pfam: PF02738; Ald_Xan_dh_C2; 1.
 DR Pfam: PF00941; FAD_binding_5; 1.
 DR Pfam: PF00111; fer2; 1.
 DR Pfam: PF01799; fer2.2; 1.
 DR ProDom: PD186071; 2Fe-2S_BD; 1.
 DR ProSite: PS00197; 2FE2S_FERREDOXIN; 1.
 DR ProSite: PS00559; MOLYBDOTERIN_EUK; 1.
 KW Oxidoreductase; NAD; Molybdenum; Flavoprotein; FAD; Iron-sulfur.
 FT INIT MET 0 0
 FT METAL 36 36 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT METAL 42 42 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT METAL 47 47 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT METAL 50 50 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 SQ SEQUENCE 1330 AA; 146111 MW; A3DD206B9D74E365 CRC64;
 Query Match
 Best Local Similarity 28.0%; Pred. No. 28;
 Matches 33; Conservative 12; Mismatches 43; Indels 30; Gaps 6;
 OY 35 TCDRHGGCGQYSA---QSRQRPQGVDXLCSAGSTVYAPFTGMTVGGEKPYQNKNA1NGV 91
 DB 660 TCVGHIIGAVVADPPEHQAARV-----KITVEDDPAITTD-----AINN- 703
 OY 92 RISRGFCVKMYKPIYV-KGPIKKGKGLTLLPLQKYVPGDLSHV 147
 DB 704 -----NSFTYGEIKIEKDKLKGSEADNVSGELIGQEHFYLETNTIAPV 752
 RESULT 10
 YTE3_CAEEL
 ID YTE3_CAEEL STANDARD; PRT; 184 AA.
 AC Q11079;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 20.9 kDa protein B0563.3 in chromosome X.
 GN B0563.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodidae; Caenorhabditis.
 ON NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Faveille T.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE BIL FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U28740; AAA68321.1; -
 DR Wormpep; B0563.3; CE02441.
 DR InterPro: IPR002199; UPF0005.
 DR Pfam: PF01027; UPF0005; 1.
 DR Hypothetical protein; Transmembrane.
 FT TRANSMEM 70 90 POTENTIAL.
 FT TRANSMEM 96 116 POTENTIAL.
 FT TRANSMEM 123 143 POTENTIAL.
 FT TRANSMEM 160 180 POTENTIAL.
 SO SEQUENCE 184 AA; 20942 MW; 09836D218202F378 CRC64;

Query Match 8.5%; Score 68; DB 1; Length 184;
 Best Local Similarity 27.2%; Pred. No. 4.2;
 Matches 31; Conservative 9; Mismatches 32; Indels 42; Gaps 7;

OY 25 CACKSNIEIRTCDRHGGQYASQ--RSQRPHQGVXLCGASG-----TVY--- 67
 Db 21 CVKRS-----RPNCSDYCAPRLKQSPSO---LCSFGRIRHSCSCHYGMLTLEAK 69
 OY 68 -----APFTGMIVGEKPY--QNKNAINGVRISGRFCV-----KMFYIKP 107
 Db 70 VVLEAAVTGTVASLFAVTLQNKRFDSVGYASMGSLLCVLLMAGIFQMFMSF 123

RESULT 11
 MUS2_AOUAE STANDARD; PRT: 762 AA.

ID MUS2_AOUAE STANDARD; PRT: 762 AA.
 AC 067287;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Muts2 protein.
 GN MUTSB OR MUTS2 OR AQ_1242.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aubay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RL Nature 392:353-358(1998).
 CC -!- FUNCTION: NOT KNOWN.
 CC -!- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

CC EMBL: AE000730; AAC07247.1; -
 DR InterPro: IPR000432; Muts_C.
 DR InterPro: IPR002863; Muts_N.
 DR Pfam: PF00488; Muts_C; 1.
 DR Pfam: PF01713; Smr; 1.
 DR ProDom: PD001263; Muts_C; 1.
 DR SMART: SM00534; Mutsac; 1.
 DR SMART: SM00533; Mutsac; 1.
 DR SMART: SM00463; Smr; 1.
 DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; FALSE_NEG.
 KW ATP-binding; DNA-binding; Complete proteome.
 FT NP_BIND 329 336 ATP (POTENTIAL).
 SO SEQUENCE 762 AA; 87538 MW; 3A7D977DC2EB01BC CRC64;

Query Match 8.5%; Score 68; DB 1; Length 762;

Best Local Similarity 21.8%; Pred. No. 19;
 Matches 27; Conservative 19; Mismatches 40; Indels 38; Gaps 4;

OY 64 STYVAPFTGMIVGEKPYQNKNAINGVRISGRF-----CVKMFYIKPIYKGIKKEK 119
 Db 20 STVSHSPATKEIKQNLKPYTNKEKVEKEIELSNFAFIAENVRLEFEEDIRL--LKKAKL 77
 OY 120 LGTLL-----PLOKYPGQSHVHEN--CDSSDP 147
 Db 78 QGALIGVEDIILKILNVILNLTKEIRVLSHVQRLEPRKRYKRLYTFSPLENIIGSIDP 137

OY 148 TAYL 151
 Db 138 RGFV 141

RESULT 12
 COL1_ARATH STANDARD; PRT: 355 AA.
 ID COL1_ARATH STANDARD; PRT: 355 AA.
 AC 050055;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Zinc finger protein CONSTANTS-LIKE 1.
 GN COL1 OR AT5G15850 OR F14F8_230.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv.
 RA Putterill J.J., Ledger S.E., Lee K., Robson F., Murphy G.,
 RA Coupland G.;
 RT "The flowering-time gene CONSTANS and homologue CONSTANS LIKE 1 exist
 RT as a tandem repeat on chromosome 5 of Arabidopsis.";
 RL (in) Plant Gene Register PGR97-077.
 RN 12

RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Nakazaki N., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Watanabe A., Yamada K., Okumura S., Shinpo S., Takeuchi C., Wada T.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoneking T., Pepl K., Spleth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Oeizsky P., Riley A., Stromwall C.,
 RA Wagner-McPherson C., Wollam A., Yoakum M., Bell M., Dedhia N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 RA Kirchoff K., Roth K., King L., Bahret A., Miller B., Marra M.,
 RA Marijnsen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
 RA Eutlan K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Roben J., Grymoprez B., Zimmermann W.,
 RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirke W., Woolman P., Klein Lankhorst R.,
 RA Weltzenegger T., Bothe G., Rose M., Hauf J., Benneiser S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villarroel R., Gietlen J., Ardiles W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K., Rudd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransz P.;
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:823-826(2000).
 RN 13
 RP CHARACTERIZATION.
 RC STRAIN=cv. C24;
 RX MEDLINE=21260187; PubMed=11359606;

RA Ledger S., Strayer C., Ashton F., Kay S.A., Putterill J.;
 RT "Analysis of the function of two circadian-regulated CONSTANS-LIKE
 RL genes.";
 CC Plant J. 26:15-22(2001).
 CC
 CC -1- FUNCTION: Putative transcription factor that may be involved in
 CC the light input to the circadian clock but does not affect
 CC flowering time.
 CC
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LEAVES AND AT LOWER LEVELS
 CC IN STEMS, FLOWERS AND SILIQUES. NOT DETECTED IN ROOTS.
 CC
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT.
 CC
 CC -1- INDUCTION: EXPRESSED WITH A CIRCADIAN RHYTHM SHOWING A PEAK
 CC AT DAWN.
 CC
 CC -1- SIMILARITY: BELONGS TO THE CONSTANS FAMILY.
 CC
 CC -1- SIMILARITY: CONTAINS 2 B BOX-TYPE ZINC FINGERS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC
 CC EMBL: Y10555; CAA71587.1; -
 CC DR EMBL: Y10556; CAA71588.1; -
 CC DR EMBL: AL391144; CAC01784.1; -
 CC DR Mendel: 24611; Arath.1362;24611.
 CC DR InterPro: IPR000315; Znf_Box.
 CC DR InterPro: IPR002926; Znf_Constans.
 CC DR Pfam: PF01760; Zf-CONSTANS; 2.
 CC DR Prodom: PD007661; Znf_Constans; 1.
 CC DR SMART: SM00336; BBOX; 1.
 CC DR PROSITE: PS01119; 2F_BBOX; 2.
 CC
 CC Transcription regulation: zinc-finger; DNA-binding; Nuclear protein;
 CC Repeat: Multigene family.
 CC
 CC ZN_FING 12 B BOX-TYPE 1.
 CC FT ZN_FING 55 B BOX-TYPE 2.
 CC FT ZN_FING 97 B BOX-TYPE 2.
 CC
 CC SEQUENCE 355 AA; 39491 MW; 412C237A9CD30309 CRC64;
 SO
 Query Match 8.4%; Score 67.5; DB 1; Length 355;
 Best Local Similarity 23.6%; Pred. No. 9.6;
 Matches 29; Conservative 15; Mismatches 44; Indels 35; Gaps 4;
 Oy 20 PWANICAGKSNRICTDRHCGCGYSAORSOPHOGVDXL-----CSAGSY 66
 Db 62 PAAFFKADAASTCTTCDSE--IHSANPLARRHQRVPLPISSEYSSTATNHSCTTV 118
 Oy 67 VAPFTGMIVGOEK-----PYONKNAINNGVIRISGRFCVKMFYIKPIYKGP 113
 Db 119 TDPENRLVIGQEEDEDEDEAEASWLLPNSGKNSGNN-----NGFSIGDEFNLVDYSSS 172
 Oy 114 IKK 116
 Db 173 DKQ 175
 RESULT 13
 VC50_HSVSA STANDARD; PRT; 535 AA.
 ID VC50_HSVSA
 AC 001012;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE Probable transcription activator EDRL1.
 GN 50 OR EDRL1.
 OS Herpesvirus saimiri (strain 11).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 CC NCBI_TaxID=10383;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92333688; PubMed=1321287;

RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
 RA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,
 RA Honess R.W.;
 RT "Primary structure of the herpesvirus saimiri genome.";
 RL J. Virol. 66:5047-5058(1992).
 CC
 CC [2]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=92230228; PubMed=1314457;
 RA Nicholas J., Coleman K.R., Coleman H., Newman C., Honess R.W.;
 RT "Analysis of the nucleotide sequence of the rightmost 43 kbp of
 RT herpesvirus saimiri (HVS) L-DNA: general conservation of genetic
 RT organization between HVS and Epstein-Barr virus.";
 RL Virology 188:296-310(1992).
 CC
 CC [3]
 CC SEQUENCE FROM N.A., AND CHARACTERIZATION.
 CC Nicholas J., Coles L.S., Newman C., Honess R.W.;
 RT "Regulation of the herpesvirus saimiri (HVS) delayed-early
 RT 110-kilodalton promoter by HVS immediate-early gene products and a
 RT homolog of the Epstein-Barr virus R trans activator.";
 RL J. Virol. 62:2457-2466(1988).
 CC
 CC -1- FUNCTION: TRANSCRIPTION ACTIVATION. REGULATES THE DELAYED-EARLY
 CC 110 kDa PROMOTER.
 CC
 CC -1- SIMILARITY: TO EBV BRLF1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC
 CC EMBL: X64346; CAA45672.1; ALT_INIT.
 CC DR EMBL: M86409; AAA46124.1; ALT_INIT.
 CC DR EMBL: M60850; AAA46159.1; ALT_INIT.
 CC DR PIR: C36811; C36811.
 CC DR HSP: P10910; ZEED.
 CC KW Transcription regulation; Activator; DNA-binding; Early protein.
 CC
 CC SEQUENCE 535 AA; 60050 MW; B4FF2B9ABA38816FB CRC64;
 SO
 Query Match 8.4%; Score 67.5; DB 1; Length 535;
 Best Local Similarity 23.0%; Pred. No. 15;
 Matches 46; Conservative 17; Mismatches 66; Indels 71; Gaps 8;
 Oy 5 KALLLAGLSTALAG-----PWANICAGKSSN--EIRTCDRHCGGYSA 46
 Db 197 QALNESCIVSTTLAALFKLSPPTISIPLEKPMFQSCGNGQNPDI-----CROGSV 249
 Oy 47 QNSORPHOGVDXCSAGSYVAPFTGMIVGOEKPYONKNAINNGVIRISGRFCVKMFYIK 106
 Db 250 IR--RPHOGV-----FEDIRPDPPLMRKISSENSFKKSTANISTLLQPKLEIDDPED 302
 Oy 107 PIKYGPIKKE-----KLGLLPLQKYVPGIO----- 134
 Db 303 PRIGGFPLNKEETATPLKDSFSNPTFINTGANTLLPRAASTPALLESFSPTHFCMSD 362
 Oy 135 -----SHVHENCSSDPT 148
 Db 363 ESTASTSHVPLDN-NISLPT 381
 RESULT 14
 TR11_MOUSE STANDARD; PRT; 625 AA.
 ID TR11_MOUSE
 AC 035305;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 11A precursor
 DE (Receptor activator of NF-kB) (osteoclast differentiation factor
 DE receptor) (ODFR).
 GN TNFRSF11A OR RANK.
 OS Mus musculus (mouse).
 RX

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal liver;
 RA MEDLINE=98032977; PubMed=9367155;
 RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C.,
 RA Tometsko M.E., Roux E.R., Teepe M.C., Dubose R.F., Cosman D.,
 RA Galibert L.;
 RT "A homologue of the TNF receptor and its ligand enhance T-cell growth
 RT and dendritic-cell function.";
 RL Nature 390:175-179(1997).
 RN [2]
 RP FUNCTION.
 RX MEDLINE=99097247; PubMed=9878548;
 RA Nakagawa N., Kinoshita M., Yamaguchi K., Shima N., Yasuda H., Yano K.,
 RA Morinaga T., Higashio K.;
 RT "RANK is the essential signaling receptor for osteoclast
 RT differentiation factor in osteoclastogenesis.";
 RL Biochem. Biophys. Res. Commun. 253:395-400(1998).
 CC -1- FUNCTION: RECEPTOR FOR RANK LIGAND (RANKL; ALSO KNOWN AS
 CC OSTEOCLAST DIFFERENTIATION FACTOR OR ODF), ESSENTIAL FOR RANKL-
 CC MEDIATED OSTEOCLASTOGENESIS. INVOLVED IN THE REGULATION OF
 CC INTERACTIONS BETWEEN T-CELLS AND DENDRITIC CELLS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: UNIFORM EXPRESSION WITH HIGH LEVELS IN
 CC TRABECULAR BONE, THYMUS, SMALL INTESTINE, LUNG, BRAIN AND KIDNEY.
 CC WEAKLY EXPRESSED IN SPLEEN AND BONE MARROW.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF019046; AAB86810.1; -;
 DR HSSP; P25942; ICDF.
 DR MGD; MGI:1314891; Tnfstfla.
 DR InterPro: IPR001368; TNFR_c6.
 DR Pfam; PF00020; TNFR_c6; 3.
 DR ProDom; PD000771; TNFR_c6; 1.
 DR SMART; SM00208; TNFR; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; 1.
 DR PROSITE; PS50050; TNFR_NGFR_2; 1.
 DR Receptor; Glycoprotein; Transmembrane; Repeat; Signal.
 FT SIGNAL 1 30
 FT CHAIN 1 625
 FT
 FT DOMAIN 31 214
 FT TRANSMEM 215 234
 FT DOMAIN 235 625
 FT REPEAT 35 69
 FT REPEAT 72 113
 FT REPEAT 115 152
 FT REPEAT 155 195
 FT DISULFID 35 47
 FT DISULFID 48 61
 FT DISULFID 51 69
 FT DISULFID 72 87
 FT DISULFID 93 113
 FT DISULFID 115 128
 FT DISULFID 134 152
 FT CARBOHYD 106 106
 FT CARBOHYD 175 175
 FT SEQUENCE 625 AA; 66621 MW; F8C1872E9511DBE CRC64;
 SO
 Query Match 8.4%; Score 67.5; DB 1; Length 625;
 Best Local Similarity 21.4%; Pred. No. 18;
 Matches 31; Conservative 11; Mismatches 40; Indels 63; Gaps 5;

QY 16 ALAGPWANI-----CAGKSNFIRCDHGGCGQYSAQSRPHQGVDXLCSA 62
 Db 309 AAGSPMAEVRDSRFTFLVSEVETQDLSKRIPEDY-----TDRSQPSTGSLILLQ 362
 QY 63 GSTVYAPFTGMVGOEKFPYQNKNAINNGVRISGRGFCVMFYIKPIKYGKIRKGEKLG 122
 Db 363 GSKSTPPE-----QE-----PLEVGENDS- 381
 QY 123 LPLQKYVPGIQSHVHIENCSSDP 147
 Db 382 ---LSQCFGTGETVDESECDPTEP 403
 RESULT 15
 MOBI_XENLA STANDARD; PRT; 398 AA.
 ID MOBI_XENLA
 AC P38565;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Integumentary mucin B.1 (FIM-B.1) (Fragment).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 RX NCBI_Taxid=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91002513; PubMed=2207068;
 RA Probst J.C., Gertzen E.-M., Hoffmann W.;
 RT "An integumentary mucin (FIM-B.1) from Xenopus laevis homologous with
 RT von Willebrand factor.";
 RL Biochemistry 29:6240-6244(1990).
 CC -1- FUNCTION: COULD BE INVOLVED IN DEFENSE AGAINST MICROBIAL
 CC INFECTIONS. PROTECTS THE EPITHELIA FROM EXTERNAL ENVIRONMENT.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: EXPRESSED AND STORED EXCLUSIVELY IN MATURE
 CC MUCOUS GLANDS OF THE SKIN.
 CC -1- PTM: EXTENSIVELY O-GLYCOSYLATED.
 CC -1- SIMILARITY: CONTAINS 1 WFMC DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; J02910; AAA49711.1; -;
 DR PIR; A35281; A35281.
 DR HSSP; P04355; AMT2.
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR001007; WFMC.
 DR Pfam; PF00007; Cys_knot; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01208; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR Repeat; Amphibian skin; Glycoprotein.
 FT NON_TER 1 93
 FT DOMAIN 1 93
 FT REPEAT <1 7
 FT REPEAT 8 18
 FT REPEAT 19 29
 FT REPEAT 30 40
 FT REPEAT 41 51
 FT REPEAT 52 62
 FT REPEAT 66 76
 FT
 8 X 11 AA TANDEM REPEATS OF G-E-S-T-P-
 A-P-S-E-T-T-
 1 (INCOMPLETE).

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2002, 16:27:23 ; Search time 28 Seconds
(without alignments)
932.937 Million cell updates/sec

Title: US-09-868-953-1

Percent score: 804

Sequence: 1 MESTKALLAGLIALAGP.....GIQSHVHTENDSDPTAYL 151

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 segs, 172994929 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

SPREMBL_19:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_prodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriaph:*

17: sp_archae:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	418.5	52.1	160	13	Q9DGE4
2	370	46.0	156	13	Q9DGE4
3	142.5	17.7	472	5	Q21241
4	87.5	10.9	689	5	001719
5	82	10.2	272	17	Q962E8
6	80.5	10.0	417	16	Q962E8
7	78	9.7	498	13	Q9PFR2
8	77.5	9.6	538	15	Q91ED5
9	76	9.5	240	17	Q9H560
10	76	9.5	651	4	Q9U893
11	74.5	9.3	711	11	Q9H879
12	74	9.2	750	16	Q98Q88
13	73.5	9.1	531	16	Q9CP23
14	73	9.1	335	12	Q91H04
15	72.5	9.0	284	2	Q9AKK4
16	72.5	9.0	428	2	Q9XK34

17	72.5	9.0	900	15	Q9QNZ8	Q9QNZ8 human immun
18	72	9.0	478	4	Q9BSK1	Q9BSK1 homo sapien
19	71.5	8.9	706	10	Q81324	Q81324 arabidopsis
20	71.5	8.9	749	10	Q9L1R8	Q9L1R8 arabidopsis
21	71.5	8.9	773	10	Q9C902	Q9C902 arabidopsis
22	71	8.8	190	12	Q9J3A1	Q9J3A1 hepatitis c
23	71	8.8	284	16	Q92GNO	Q92GNO rickettsia
24	70.5	8.8	779	17	Q97A76	Q97A76 thermoplasma
25	70.5	8.8	548	15	Q91ED6	Q91ED6 human immun
26	70.5	8.8	673	4	Q9HC21	Q9HC21 homo sapien
27	70.5	8.8	680	4	Q9NVW4	Q9NVW4 homo sapien
28	70	8.7	352	2	Q9AHR2	Q9AHR2 uncultured
29	70	8.7	444	10	Q9SUP5	Q9SUP5 arabidopsis
30	70	8.7	644	17	Q974G9	Q974G9 sulfolobus
31	69.5	8.6	263	11	Q9R1O6	Q9R1O6 mus musculus
32	69.5	8.6	263	11	Q9WU52	Q9WU52 mus musculus
33	69.5	8.6	268	16	Q97DM2	Q97DM2 clostridium
34	69.5	8.6	472	4	Q96PE6	Q96PE6 homo sapien
35	69.5	8.6	715	10	Q9LND9	Q9LND9 arabidopsis
36	69	8.6	186	12	Q9J3D9	Q9J3D9 hepatitis c
37	69	8.6	191	12	Q9O3X9	Q9O3X9 hepatitis c
38	69	8.6	1262	5	Q95VG2	Q95VG2 caenorhabdi
39	69	8.6	1418	5	Q44518	Q44518 caenorhabdi
40	68.5	8.5	734	2	Q85122	Q85122 shewanella
41	68.5	8.5	735	2	Q87539	Q87539 shewanella
42	68	8.5	171	12	Q91IT4	Q91IT4 hepatitis c
43	68	8.5	186	12	Q9J3E0	Q9J3E0 hepatitis c
44	68	8.5	186	12	Q9J3D8	Q9J3D8 hepatitis c
45	68	8.5	186	12	Q9J3D7	Q9J3D7 hepatitis c

ALIGNMENTS

RESULT 1

ID Q9DGE4 PRELIMINARY; PRT; 160 AA.

AC Q9DGE4;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE MYELOID PROTEIN-1.

OS Cyprinus carpio (Common carp).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Cyprinus.

OX NCBI_TaxID=7962;

RN [1]

RP SEQUENCE FROM N.A.

RA Fujiki K., Nakao M., Shin D., Yano T.;

RT "Molecular cloning of a carp homolog of chicken myeloid protein-1.";

RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.

DR EMBL; AB027192; BAB16024.1; ;

SQ SEQUENCE 160 AA; 17461 MW; BD15A9A910B4324C CRC64;

Query Match 52.1%; Score 418.5; DB 13; Length 160;

Best Local Similarity 50.3%; Pred. No. 1.5e-38;

Matches 78; Conservative 28; Mismatches 36; Indels 13; Gaps 4;

QY 6 ALLAGLIALA-----GPMANICAGKSSNEITRCD-RHGGQYSNOR--SQRPHG 55

DB 8 SFLILAVICSLVDASQVKEFP--LCSGNSNRRKGCCKKYCGNGVSGSRDGGKRRHAG 64

QY 56 VDLCSGTVVYAPFTGMIYGOECPYQKNATNNGVRISGRGCVKMFYIKPIKYGPIK 115

DB 65 LDIIVCADGATVYAPFDVKNLNGKAVPYKNNALINDSINSGGGLCFKLFYVPISTYGTGK 124

QY 116 KGEKLTLPLOKVPYGIQSHVHTENDSDPTAY 150

DB 125 KGGKIGTMLPMQKYYPGITSHVHVQMCDBRSDPTKY 159

RESULT 2

```

ID 09DFJ1
AC 09DFJ1 PRELIMINARY; PRT; 156 AA.
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CHEMOTAXIN (FRAGMENT).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NBI_TaxId=8022;
RN [1]
RP
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER.
RA Bayne C.J., Gerwick L., Fujiki K., Nakao M., Yano T.;
RT "Immune-related (including acute phase) genes identified in the
RT livers of rainbow trout, Oncorhynchus mykiss, by means of suppression
RT subtractive hybridization."
RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
EMBL: AF271114; AAC28030.1; -.
SQ
SQ NON_TER 156
SQ SEQUENCE 156 AA; 17087 MW; B19067665FDA70DA CRC64;

```

Query Match	46.0%;	Score 370;	DB 13;	Length 156;
Best Local Similarity	45.6%;	Pred. No. 3,6e-33;		
Matches	68;	Conservative	29;	Mismatches 47; Indels 2; Gaps 1
QY	7	LLAGLISLALAGPWANICAGKSSNEIRTCDRHGSGQGYSAQFSQRPHQGVDLCSAGSTV	66	
	::::	::::	::::	::::
Db	10	VLLAVLSECEWEVKKRGQGLSSNNRRRTGDRMGQGHGDCSNGRANGLDLYCNDGATV	69	
QY	67	YAPFTGMTVGGDEKPYQN--KNAIINGVRIISGEGFCVKKMEYIIPRIKYKGIKKGEKGIITLL	124	
		::::	::::	::::
Db	70	YAPFVTKINGKVIYVTDKKAALINDGINLSGGLCEFKLFYVPRDKYSGVYKKGQRTGLM	129	
QY	125	PLQKRYPGIQSHVHTIENCDSSDPYAT	150	
	::::			
Db	130	TMQSVYPGITSHVHVMOCDSKSDPTKF	155	

```

RESULT 3
ID 021241
AC 021241; PRELIMINARY; PRT; 472 AA.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DT HYPOTHETICAL 53.3 KDA PROTEIN.
DT K05F1.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Wohldmann P.;
RT "The sequence of C. elegans cosmid K05F1."
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submissiion."

```

[illegible]

RESULT 4					
ID	001719	PRELIMINARY:	PRT:	689 AA.	
AC	001719;				
DT	01-JUL-1997 (TREMBLrel. 04, Created)				
DT	01-JUL-1997 (TREMBLrel. 04, Last sequence update)				
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)				
DE	MYB-RELATED TRANSCRIPTION FACTOR (FRAGMENT).				
OS	Strongylocentrotus purpuratus (Purple sea urchin).				
OC	Eumetazoa; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;				
CC	Echinoidae; Euechinoidae; Echinacea; Echinoida; Strongylocentrotidae;				
OX	NCR1_TaxID=7668;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96088684; PubMed=9428408;				
RA	Cotifman J.A., Kirchhamer C.V., Harrington M.G., Davidson E.H.;				
RT	"SmYd functions as an intramodular repressor to regulate spatial				
RT	expression of Cyfilla in sea urchin embryos.";				
RL	Development 124:4717-4727(1997).				
DR	EMBL; U96090; AAC47807.1; --.				
DR	HSSP; P06876; IMBK.				
DR	InterPro; IPR001005; MYD_DNA_bind.				
DR	Pfam; PF00249; myb_DNA-binding; 3.				
DR	SMART; SMO0395; SANT; 3.				
DR	PROSITE; PS00037; MYB_1; UNKNOWN_3.				
DR	PROSITE; PS00334; MYB_2; 3.				
DR	PROSITE; PS50090; MYB_3; 3.				
FT	NON_TER				
SQ	SEQUENCE 689 AA; 77241 MW; 0FA5A8D7AF8766C6 CRC64;				
<hr/>					
Query Match	10.9%; Score 87.5; DB 5; Length 689;				
Best Local Similarity	22.4%; Pred No. 0.45;				
Matches 36; Conservative	22; Mismatches 46; Indels 57; Gaps				
QY	17 LAGPANIC--AGKSSNEIR-----TCDR--HGGGQYSNORSQR-PH---QGVDXLCS 61				
Dbb	163 LGNRKAELAKLLPGTLDNAIKRWMSSTMKRYETCNPTPRSTTKRPHYTHINDNKPCS 222				
QY	62 AGSTVAPFPTGMIVGOEKPYONKNALINNGCVRLSGRGFCVKMPLYIKPIKTIGPKKGKGLG 121				
Ddb	223 SSKSIYTP-----DSDFMNANSIRDLAIRMGOGQRFV-----VR 255				
QY	122 TLLEPL-----QKVYPGI---QSHVIENDSSDP 147				
Ddb	236 TLYPMGHDTLGODEGEGRSVKPGVATPKQWMLINDNGEITSP 296				
<hr/>					
RESULT 5					
ID	096ZEB	PRELIMINARY:	PRT:	272 AA.	

```

AC 0962E8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHEMETICAL PROTEIN ST1885.
GN ST1885.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=11955;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001)
DR EMBL; AP000988; BAB65977.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 272 AA; 30871 MW; 810C9741373BF05 CRC64;

Query Match 10.2%; Score 82; DB 17; Length 272;
Best Local Similarity 29.7%; Pred. No. 0.59; Mismatches 34; Indels 12; Gaps 5;
Matches 27; Conservative 18;

QY 67 YAPFTGMI-----VGQEPYONKNAINGVRISGRGCVKMFYKPIKYG-PIKGE 118
DB 36 YSPFGIIEKIKRKIGRPKY-TKIDYDVVMYNSNGKRITKVLHVEPYLESGSEIKGE 94
QY 119 KLGTLLPQKV-PEIQSHHINCSDSP 148
DB 95 KIGKE--LESPTAGDFKHAHIGITFKFPS 123

RESULT 6
Q9PGX9 PRELIMINARY; PRT; 417 AA.
AC Q9PGX9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 15, Last annotation update)
DE HYPOTHEMETICAL PROTEIN XF0167.
GN XF0167.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OX Xylella.
OX NCBI_TaxID=2371;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arraya P., Abreu F.A., Acencio M.,
RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitaajima J.P.,
RA Klieger J.E., Kuramae E.E., Laligret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,

```

```

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quagga R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Zeldin J., Zebubai J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AE003870; AAF82980.1;
DR MEROPS; M37.DPW;
DR InterPro; IPR002886; Peptidase_M37.
DR Pfam; PF01551; Peptidase_M37; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 417 AA; 46554 MW; 5483743CD61061FD CRC64;

Query Match 10.0%; Score 80.5; DB 16; Length 417;
Best Local Similarity 25.0%; Pred. No. 1.5; Mismatches 42; Indels 21; Gaps 7;
Matches 29; Conservative 24;

QY 34 RTCDRHGCGOYSAORSQRPQGVDXLCSAGSTVYVAPFTGMI--VGQEPYONKNAINGV 91
DB 249 RLSTFGLRRHPILGTRMRHKGVYAAATGPIMAGDARVOFGQGRGVNVLNH-- 306
QY 92 RLSRGECYKMFYKPIKYGKPKKGL--CTLLPQKVPGIOS----VHTE 140
DB 307 ---GKGY--RTLGHMSRF-GKIKAGKINGTVIG----YVGMGLATGPHLHYE 352

RESULT 7
Q9PTR2 PRELIMINARY; PRT; 498 AA.
AC Q9PTR2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CYTOCHROME P450 2P2.
GN CYP2P2.
OS Fundulus heteroclitus (Killifish) (Mummichog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Fundulidae; Fundulus.
OX NCBI_TaxID=8078;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Oleksiak M.F., Stegeman J.J.;
RT "Diversity of vertebrate cytochrome P450 2 family genes: Cloning of
RT three new genes and phylogenetic analysis of CYP2 gene subfamilies in
RT Fundulus heteroclitus.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN 12
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Oleksiak M.F., Wu S., Parker C., Zeldin D.C., Stegeman J.J.;
RT "Cloning and expression of a novel teleost cytochrome P450, CYP2P3;
RT conservation of arachidonic acid epoxidase/19-Hydroxylase.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF117342; AAF21999.1;
DR HSSP; P00179; 1D76.
DR InterPro; IPR001128; Cyt_P450.
DR Pfam; PF00067; P450; 1.
DR PROSITE; PS00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 498 AA; 57340 MW; 1FF4ED84B8D610C9 CRC64;

Query Match 9.7%; Score 78; DB 13; Length 498;
Best Local Similarity 21.7%; Pred. No. 3.4;

```

Matches 40; Conservative 28; Mismatches 58; Indels 58; Gaps 10;

QY 1 MESTFALLAGLIALAPW-----ANICAGKSSNEIRTCORHGGCGYSAQS 49
 DB 117 VFKNGIVMSN-----GYPMKYORRALHALRFGGLGKKTME-----RTIQQC 160
 QY 50 Q-----RPHGVXILCSAGSTVYAPFTGMIVGOEKPYONK-----NAINNGVR 92
 DB 161 QYLNVEFYDQCKPFGSGQLINNANSNI-----ICLVFNGNREYDCKEHTLIDNNELLR 217
 QY 93 TSGRGFCVAKMYIKPIKYR---GPIKKGKGLTLLPLQKYPGIOSHV--HIENCDSDP 147
 DB 218 LQG-GFWQVYVMFSPVMKWLPGPHKK-----IFIHLQKIIDFLFIRIKHRENDLPSSP 271
 QY 148 TAYL 151
 DB 272 RDIY 275

SUBT 8
 ED5

QYIED5 PRELIMINARY; PRT; 538 AA.

AC QYIED5; PRELIMINARY; PRT; 538 AA.
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE ENV POLYPEPTIDE (FRAGMENT).
 CN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 ON NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-YBF51.
 RA Roques P., Robertson D., Diamond F., Sousquiere S., Mauciere P.,
 RA Dejeune C., Brun-Vezinet F., Dormont D.,
 RT "HIV-1 group O phylogenetic analysis of C2-gp41 region."
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ133072; CAB96233.1; -
 DR InterPro: IPR000328; Env.GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.
 KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.
 FT NON_TER 1
 FT NON_TER 538
 SQ SEQUENCE 538 AA; 60777 MW; B3C9E6A233FEFID CRC64;

Query Match 9.6%; Score 77.5; DB 15; Length 538;
 Best Local Similarity 26.9%; Pred. No. 4.3;
 Matches 36; Conservative 13; Mismatches 42; Indels 43; Gaps 6;

QY 25 CAGKSSN--EIRTCDRHGGCGYSAQSRPHGVXILCSAGSTVYAP-----PT 71
 DB 175 CEGKSCNFKENRT-----NNNTQIPCKLKQVRSMMKGGSLTAPLRGNLTCSNIT 227
 QY 72 GMTVGOEKPYONKAINNGVRISG-----RGFCVKMFYIKPIKYGPI--- 114
 DB 228 GWTLDLDKRW-NRNDNTNFRPFGGKMKDLWRTELLRYKVKFVSAPTRLARPVIS 286
 QY 115 -----KKGKGLTL 123
 DB 287 GTHREKRAVGLGWL 300

RESULT 9

ID Q9HS60 PRELIMINARY; PRT; 240 AA.

AC Q9HS60;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
 DE VNG0391C.

VNG0391C.

OS Halobacterium sp. (strain NRC-1).

OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;

OC Halobacterium

ON NCBI_TaxID=64091;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20504483; PubMed=11016950;

RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,

RA Shukla H.D., Lasky S.R., Balla N.S., Thorsson V., Sirogna J.,

RA Swartzell S., Welt D., Hall J., Dahl T.A., Weltl R., Goo Y.A.,

RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,

RA Madocks D.G., Jablonski P.E., Krebs M.P., Angewine C.M., Dale H.,

RA Isenbarger T.A., Beck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,

RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,

RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,

RT "Genome sequence of Halobacterium species NRC-1";

RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

DR EMBL: AE004996; AAG18948.1; -

DR InterPro: IPR002761; DUF71.

DR Pfam: PF01902; DUF71; 1.

KW Complete proteome.

SQ SEQUENCE 240 AA; 25855 MW; FCAE7C983ECEF72A CRC64;

QY 9 LAGLSTALAGPWANICACKSSNEIR-----CDRHGGCGYSAQSRPHGVXILCSA 62
 DB 98 VADTLDGGLAG-----VIAGVESDFQSRMGWCRLGCELYAPIMQPPRELADMLTA 153

QY 63 GSTVYAPFTGMIV-----GOEKPYONKN-----AINN--GVRSIGRCVYKMYI 105
 DB 154 G-----FEIRLIQVAYAGDKSMGLGRRLDRLDALALALANDEGVHMLGEGEFETLVY 207

QY 106 KPIKKGPIKKGKGLTLLPLQKYVPGIOSVH-ENC 142
 DB 208 DAPHMDRPIR-----LDYETVWAGDRGHIEITEAC 237

RESULT 10
 ID Q9UP93 PRELIMINARY; PRT; 651 AA.
 AC Q9UP93;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
 DE HYPOTHETICAL 70.4 KDA PROTEIN (FRAGMENT).
 GN DKF2P566N1047.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-KIDNEY;
 RA Ottenwaelder B., Obermaier B., Mewes H.W., Gaassenhuber J., Wiemann S.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL133109; CAB61414.1; -
 DR InterPro: IPR002965; P_rich_extensn.
 DR InterPro: IPR001374; R3H.
 DR Pfam: PF01424; R3H; 1.
 DR PRINTS; PR01217; PRICEXTENS.
 DR SMART; SM00393; R3H; 1.
 KW Hypothetical protein.
 FT NON_TER 1
 FT NON_TER 651
 SQ SEQUENCE 651 AA; 70424 MW; 4A8DF18AB50AD04C CRC64;

Query Match 9.5%; Score 76; DB 4; Length 651;
 Best Local Similarity 31.4%; Pred. No. 7.9;
 Matches 27; Conservative 14; Mismatches 25; Indels 20; Gaps 5;

QY 43 QYSAQRSPRHGVXLCASGTVAP-----FTGMIVGQEPYONKNAINN--GVRI 93
 DB 500 QYNAQRSGQQMPQ-----AAQAGYQPVLSGQGFGLI-GVQQRPSQSNVYINQGGTPV 552
 QY 94 SGRGFCVAKFYIKPIKYKPIKKGK 119
 DB 553 QS-----VWVSTPTMSSYQVPMTOGSO 574

RESULT 11

088679 PRELIMINARY; PRT; 711 AA.
 AC 088679;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE TRIP PROTEIN (FRAGMENT).
 GN TRIP.
 OS Cricetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetus.
 OX NCBI_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98337902; PubMed=9671805;
 RA Wilson S.A., Brown E.C., Kingsman A.J., Kingsman S.M.;
 RT "TRIP: a novel double stranded RNA binding protein which interacts
 with the leucine rich repeat of Flightless I";
 RL Nucleic Acids Res. 26:3460-3467 (1998).
 DR EMBL: AJ223076; CAA11077.1; .
 FT NON-TER
 SQ SEQUENCE 711 AA; 77173 MW; 4573468DCCD67991 CRC64;

Query Match 9.3%; Score 74.5; DB 11; Length 711;
 Best Local Similarity 24.8%; Pred. No. 13;
 Matches 34; Conservative 22; Mismatches 66; Indels 15; Gaps 5;

QY 20 PWANICAGKSSNEITRTCDRHGCGQYSAQRSPRHGVXLCASGTVAP-----PPTGM 73
 DB 548 PEOKITAG--SSEHVECTLEDENEFNGKQNGCEEDGVNTRAGDGDPLMCEADAQSVSTS 605
 QY 74 IVGQERYONKNAINNVR--ISGRGFCVAKFYIKPIKYKPIKKGKLTLLPLQVYP 131
 DB 606 ABESEETERRADDPAGGGRVDLDONSPOCKEETISPMKKRKPSSECHVL---SQHP 662
 QY 132 GIQSHVHIENC--DSSD 146
 DB 663 GQMEKKAIDGCSIDNSD 679

RESULT 12

098088 PRELIMINARY; PRT; 750 AA.
 AC 098088;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE LIPOPROTEIN.
 GN MYPV_4780.
 OS Mycoplasma pulmonis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=UAB CTIP;
 RX MEDLINE=21267165; PubMed=11353084;
 RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
 RA Mosser I., Dydyig K., Wroblewski H., Viari A., Rocha E.P.C.,
 RA Blanchard A.;
 RT "The complete genome sequence of the murine respiratory pathogen
 Mycoplasma pulmonis.";

RL Nucleic Acids Res. 29:2145-2153 (2001).
 DR EMBL: AL445564; CAC13651.1; .
 DR MYPULIST; MYPV_4780; .
 KW Complete Proteome.
 SQ SEQUENCE 750 AA; 86325 MW; 37FF06FE161602D5 CRC64;

Query Match 9.2%; Score 74; DB 16; Length 750;
 Best Local Similarity 22.3%; Pred. No. 16;
 Matches 41; Conservative 25; Mismatches 56; Indels 62; Gaps 8;

QY 12 LISTALGPWANICAGKSSNEITRTCDR--HGCGQYSAQRSPRHGVXLCASGTVAP 69
 DB 456 LQSPVSGGWLID-----DRSTVSTSKASDKNKKDHLGEDIILQNKNEVIAP 502
 QY 70 FTGMIVGQ---EKPYQ-----NKNAINNGVRISSGRGFCVAKMP-- 103
 DB 503 FDGKITIASYVAPSPYQAVGLGVITVLEVMKKDLVGQIDQSVYDQQLAETDITVYIAFMILN 562
 QY 104 -----YIKPIKYKG-----PIKKGKLTLLPLQVYPGIQSHVITE-NCD 143
 DB 563 PSYLENVKLVLEVQSSSTAIEITPATPKYKKEVIG-LVGEFKNNGGMPHVLHLEVSIG 621
 QY 144 SSDP 147
 DB 622 STNP 625

RESULT 13

09CP23 PRELIMINARY; PRT; 531 AA.
 ID 09CP23;
 AC 09CP23;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL PROTEIN PM0243.
 GN PM0243.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Pasteur M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida PM70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).
 DR EMBL: AE006058; AAK02327.1; .
 DR MEROPS; M37.0PW; .
 DR InterPro; IPR002086; Aldehyde-dehydr.
 DR InterPro; IPR002886; Peptidase-M37.
 DR Pfam; PF01551; Peptidase_M37; 1.
 DR PROSITE; PS00687; ALDEHYDE-DEHYDR-GLU; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 531 AA; 59590 MW; 93617D425C09F47D CRC64;

Query Match 9.1%; Score 73.5; DB 16; Length 531;
 Best Local Similarity 23.4%; Pred. No. 12;
 Matches 30; Conservative 18; Mismatches 47; Indels 33; Gaps 4;

QY 24 ICAGKSSNEITRTCDRH-----GCGQYSAQRSPR-----PHQ 54
 DB 346 ISAGKSYVAIQASNGRYVNRQGETLGKGFARVPLDQAVSSHFNFRPVRGRIAPHK 405
 QY 55 GVDXLCASGTVAPFTGMIVGQEPYONKNAINNNGVRISSGRGFCVAKMP--IKYKG 112
 DB 406 GVDPAFVGPVLPADGIV--EKVAYQANGAGRIYVNRHGREYQTVYMHLSRALYKAGQ 463
 QY 113 PIKKGKLT 120
 DB 464 NWKRGQRI 471

RESULT 14

091HQ4

PRELIMINARY: PRT: 335 AA.

ID 091HQ4

AC 091HQ4

DT 01-DEC-2001 (TEMBLrel. 19, Created)

DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)

DE POLYPROTEIN (FRAGMENT).

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=11103;

RN [1]

RP SEQUENCE FROM N.A.

RA Qian C., Zhi-Meng L., Ming Y.;

RT "Variant Analysis and Immunogenecity Prediction of Envelope Gene of

HCV strains from China."

Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RL EMBL; AF403236; AAK95635.1; -

DR

NON_TER

NON_TER

SEQUENCE

335 AA; 36215 MW; A3155D5CE2CDEDE1 CRC64;

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

Db 226 KYAHKEMSYKEGKIKRGO-----FIGIOGSTGNATGEHLHFE 264

Search completed: October 21, 2002, 16:33:56
Job time : 32 secs

Query Match 9.0%; Score 72.5; DB 2; Length 284;
 Best Local Similarity 23.6%; Pred. No. 7.1;
 Matches 26; Conservative 19; Mismatches 36; Indels 29; Gaps 6;

QY 44 YSAQRSGRP-HQGVNXLCSAGSTVYAPFTGMTV--GQEKPYONKNAIINNGVRISGRGFCV 100

DB 171 YKRRKKKCPHSGIDLOAKKAPITAAASGIVIKARAPDYGPFVELKH-----GRKFVT 225

Y 101 KMFYIKPIKYK--GPICKGEKLTLLPLQKVPYGIQ-----SHVHIE 140

DB 265 LASCRIPTDKFAOG--WCPITTYDAPDSSDQKPY 294

RESULT 15

09AKK4

PRELIMINARY: PRT: 284 AA.

ID 09AKK4

AC 09AKK4

DT 01-JUN-2001 (TEMBLrel. 17, Created)

DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)

DE HYPOTHETICAL 32.4 KDA PROTEIN.

OS Rickettsia montana.

OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

OC Rickettsiaceae; Rickettsia.

OX NCBI_TaxID=33991;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21219194; Pubmed=11319266;

RA Andersson J.O., Andersson S.G.E.;

RT "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes";

RL MOL. Biol. Evol. 18:829-839 (2001).

DR EMBL; AJ293330; CAC33652.1; -

DR InterPro; IPR002886; Peptidase_M37.

DR Pfam; PF01551; Peptidase_M37; 1.

KW Hypothetical protein.

KW SEQUENCE 284 AA; 32423 MW; 6708E9F3A54EF1F5 CRC64;

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE

SEQUENCE